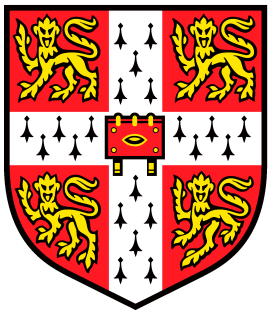
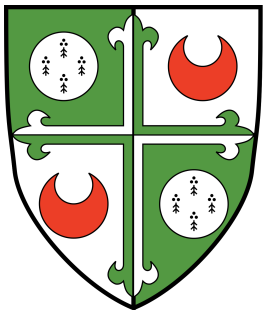

UNIVERSITY OF CAMBRIDGE

INSTITUTE OF ASTRONOMY

QUASARS AT THE HIGH REDSHIFT FRONTIER

SARAH ELENA IVANA BOSMAN

GIRTON COLLEGE

UNDER THE SUPERVISION OF

PROF MARTIN HAEHNELT

PROF GEORGE BECKER

A DISSERTATION SUBMITTED TO THE
UNIVERSITY OF CAMBRIDGE FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

Submitted to the Board of Graduate Studies
October 20, 2017

Success consists of going from failure to failure without loss of enthusiasm.

– Winston Churchill

DECLARATION OF ORIGINALITY

I, Sarah Elena Ivana Bosman, declare that this Thesis entitled *'Quasars at the High Redshift Frontier',* and the work presented in it, are my own. I confirm that this work was done wholly while in candidature for a research degree at this University, and that this Thesis has not previously been submitted for a degree or any other qualification at this University or any other institution. This Thesis is the result of my own research and contains nothing resulting from collaboration, except where explicitly noted. Where I have consulted the published work of others, this is always clearly attributed, especially in Chapters I and II which serve as an introduction and largely draw on work in the literature. I am indebted to several people, credited in the Acknowledgements, for providing comments on drafts of this Thesis at various stages of its history. The length of this Thesis does not exceed the stated limit of the Degree Committee of Physics and Chemistry of 60,000 words.
The following parts of this Thesis have already been published in, or submitted to, academic journals:

- **Chapter III** is based on the published work
  'Re-examining the case for neutral gas near the redshift 7 quasar ULAS J1120+0641'
  Bosman S. E. I. & Becker G. D., 2015, MNRAS, 452, 1105

- **Chapter IV** is based on the published work
  'A deep search for metals near redshift 7: the line-of-sight towards ULAS J1120+0641'
  Bosman S. E. I., Becker G. D., Haehnelt M. G., Hewett P. C., McMahon R. G., Mortlock D. J., Simpson C. & Venemans B. P., 2017, MNRAS, 470, 1919

- **Section VI.2.2** is adapted from the manuscript
  'A tale of seven narrow spikes and a long trough: constraining the timing of the percolation of HII bubbles at the tail-end of reionization with ULAS J1120+0641'
  Chardin J., Haehnelt M. G., **Bosman S. E. I.** & Puchwein E., MNRAS, submitted

- **Section VI.2.4** is adapted from the the submitted proposal
  'Probing galaxy - IGM interactions near reionisation with MUSE'
  P.I. **Bosman S. E. I.**, ESO Period 100A

_____

Signed

_____

Date

ACKNOWLEDGEMENTS

I thank my supervisors Martin Haehnelt and George Becker for correcting my terrible writing and other invaluable advice over the years.

I thank the reviewers, Bob Carswell and Stefano Cristiani, for helping improve this thesis and a pleasant viva.

I thank the staff of the University of Cambridge and Girton College for providing a brilliant environment in which I spent way too much time. In particular, I thank my undergraduate supervisors and apologise for all the times I turned up late and didn't hand work in. I've been on the other side now. I know your pain.

I thank the Institute of Astronomy for providing coffee, tea and biscuits that I probably still owe money for, and for being all-around awesome.

I thank Alston Misquitta, Ioana, Mark and Maria for welcoming me to Cambridge and bombarding me with difficult puzzles most of which I never solved, and delicious food most of which I couldn't finish.

I thank fellow PhD students Tricia Larsen, Kaloian Lozanov and Bjoern Soergel for explaining cosmology to me like I'm five. Our office was the best. Special thanks to Frank.

I thank Nick Woods for providing constant distraction over upwards of 700,000 online messages.

I thank Judd Harrison, with whom I came up with many great ideas which never actually worked... yet.

I thank Nicolas Dupré for insisting I should go out more.

I thank my cohort: Lindsay Oldham, Matthew Smith, Clare Wethers, Matt Read, Dominika Borucka, Bjoern Soergel, Nimisha Kulkarni, Deyan Mihailov, and Anna Shooneveld for being the best group of friends I've ever had. Seriously. You guys are the best.

I thank Marco, Alcor and Mizar for providing a great, fluffy environment in which this thesis could be finished.

I thank my parents for insisting I could make it, and my siblings Axelle and Maxime for being adorable.

Sarah Bosman

Cambridge, October 20, 2017

QUASARS AT THE HIGH REDSHIFT FRONTIER

SARAH ELENA IVANA BOSMAN

SUMMARY

In recent years the formation of primordial galaxies, cosmic metal enrichment, and hydrogen reionisation have been studied using both refined observations and powerful numerical simulations. High-redshift quasars have become a ubiquitous tool in the study of this era with more than 115 quasars now spectroscopically confirmed at $z > 6.0$. In this thesis, I use spectra of high-redshift quasars to provide improved observational constraints through a mixture of existing and new techniques.

I first investigate the claim of neutral gas around the most distant known quasar, ULAS J1120+0641(J1120), with a cosmological redshift of $z = 7.1$. Its spectrum shows a relatively weak Lyman-α emission line, which has been interpreted as evidence of absorption by neutral gas. Attributing this to a Gunn-Peterson damping wing has led to claims that the intergalactic medium is at least 10% neutral at that redshift. However, these claims rely on a reconstruction of the unabsorbed quasar emission. Initial attempts using composite spectra of lower-redshift quasars mismatched the C IV emission line of J1120, a feature known to correlate with Lyman-α and which is strongly blueshifted in J1120. I attempt to establish whether this mismatch could explain the apparently weak Lyman-α emission line. Carefully selecting quasars from SDSS DR4 based on the similarity of their C IV emission to that of J1120, I find that among a C IV-matched sample the Lyman-α line of J1120 is not anomalous. This raises doubts as to the interpretation of absorbed Lyman-α emission lines in the context of reionisation.

I then use a high quality X-Shooter spectrum of the same $z = 7$ quasar to measure the abundances of diffuse metals within one billion years of the Big Bang. I measure the occurrence rates of C IV, C II, Si II, Fe II and Mg II, producing the first measurement at $z > 6$ for many of these ions. I find that the incidence of C IV systems is consistent with a continuing decline in the total mass density of highly ionized metals, a trend seen at lower redshifts. The ratio C II /C IV, however, seems to remain constant or increase with redshift, in line with predictions from models which include a decline of the ionising ultraviolet background. The evolution in Mg II appears somewhat more complex; while the number density of strong systems continues to decline at high redshift,the number density of weak systems remains high and may even increase. This could signal an increase with redshift in the cross-section of low-ionisation metals, similar that seen previously for O I. Large numbers of weak Mg II systems are also seen at $z \sim 2$, suggesting they were already largely in place when reionisation was ending.

I use the same X-Shooter spectrum to study metal absorbers associated with the $z = 7$ quasar itself. I find that one such absorber shows signs of only partially covering the line of sight, and investigate the possible implications for the quasar's environment.

Finally, I investigate the evolution of the intergalactic medium's Lyman-α opacity using spectra of quasars at $5.7 < z < 7.1$. I assemble a sample of 92 quasar spectra, more than 3 times larger than previous samples. This allows the variance in transmission between lines of sight to be robustly quantified. The sample consists of quasars drawn from DES-VHS, SDSS and SHELLQs, as well as new reductions of archival data, and new data taken on the ESI spectrograph. I develop methods to quantify the opacity distribution, providing measurements of the distribution function up to $z = 6.1$. I find that the Lyman-α opacity evolves strongly with redshift. The scatter may be even larger than previously appreciated, posing a serious challenge for models of reionisation.

Contents

List of Figures

List of Tables

I

The First Billion Years

As the Universe expands, it gets cooler and cooler.

Neil deGrasse Tyson

Abstract

In this Chapter, I provide an overview of some aspects of our current knowledge on the Early Universe that are particularly relevant to the research presented in this thesis. The first billion years of the Universe sees the foundation of the Cosmic Web, the formation of the first stars and the assembly of the first galaxies. I review these processes in turn, before turning to the process which rendered the Universe transparent to the UV emission from stars and accreting black holes: reionisation.

I.1 Introduction

The first billion years of the Universe saw the formation of the first stars, black holes and galaxies. During this formative period, the Universe transformed from the nearly featureless mix of dark matter and primordial elements left behind by inflation and reheating, into an environment we would recognise as the cosmos today. The physical processes controlling this transition are remarkably well understood. As we push the observational frontier earlier and earlier into the history of the Universe, we begin uncovering the first galaxies to ever populate it: the current record holder is a line-emitting galaxy observed a mere 400,000,000 years after the Big Bang. Meanwhile, cosmological simulations of the early Universe have become increasingly realistic as more and more relevant physical processes have been incorporated into them. The field has moved from gravitation-only simulations of the Cosmic Web to radiative transfer models

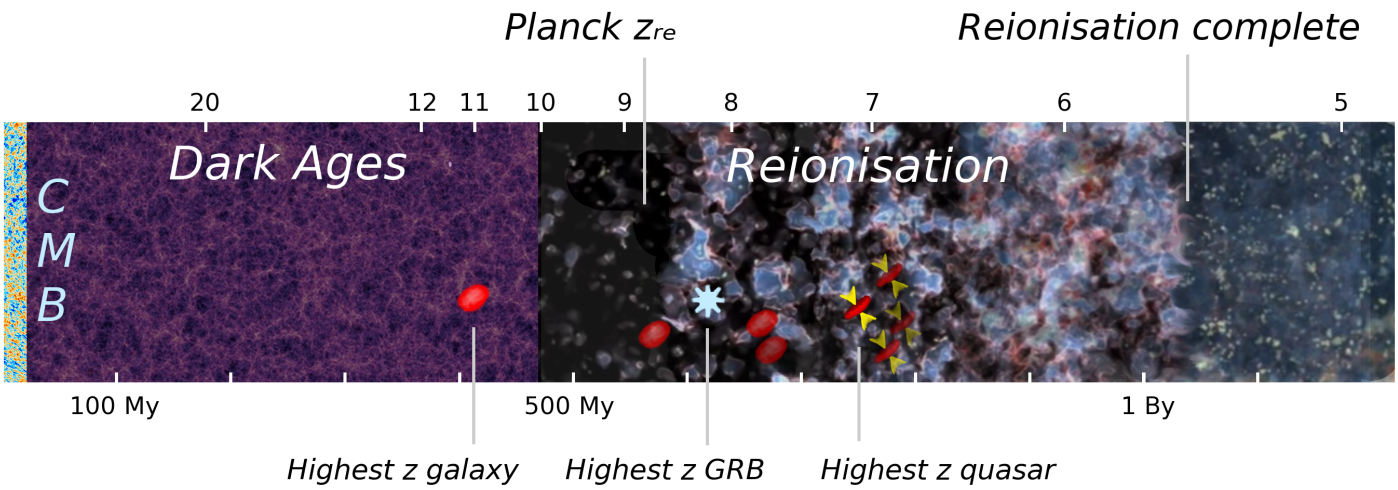

Figure I.1 – Timeline of the first billion years. Background images are a composite from Alvarez et al. (2009) and Springel (2005).

including the spectral profiles of primordial objects, temperature, and mean free path effects in the description of baryonic matter physics.

The combination of these two fronts, observational and numerical, has yielded impressive progress. For the first time, a coherent picture of the first billion years has emerged, which while not yet complete, consistently encompasses our present knowledge and lays the path to future refinements. In this Chapter, I will review important aspects of this model and outline the supporting observational evidence. I will proceed in chronological order: first, describing our understanding of cosmological structure formation, the Λ-CMD paradigm, and the formation and evolution of the Cosmic Web. I will then discuss the assembly and evolution of the first galaxies within this framework, including the current mysteries surrounding the early production and dispersion of metals. Finally, the process of reionisation, through which the neutral hydrogen filling the Universe separates into protons and free electrons, will be outlined, as will the competing models and challenges surrounding what is often called the last universal phase transition.

I.2 Structure Formation

By the end of the first 4 minutes or so after the Big Bang, the primordial high-energy physics responsible for the origin of baryonic and dark matter had ended. In the current view, the process of inflation had stretched space exponentially during the first $\sim 10^{-32}$ s, giving rise to the locally homogeneous matter distribution and geometrically flat space-time which we see today. Reheating and nucleosynthesis populated space with a baryonic mix consisting of 75% hydrogen, 24% helium and $< 1\%$ heavier elements. Quantum fluctuations arising spontaneously got frozen out by rapid expansion and became regions of space with slight over- or under-densities, which will serve as the seeds of the first structure: the Cosmic Web.

Currently our best model of cosmology is the Λ-CDM paradigm, which describes the Universe as a topologically flat Friedman-Walker metric with the addition of the cosmological constant Λ at late times, often re-branded as 'Dark Energy'. The energy density of the present-

day Universe is divided into 69.1% of this Dark Energy (Frieman et al., 2008), 4.9% baryonic matter, and 26.0% cold dark matter (CDM), with these ratios varying across cosmic time (for more details, refer to Weinberg 2008). Throughout this thesis, we will use the cosmological parameters most recently measured by the Planck satellite: $[\Omega_m, \Omega_\Lambda, h] = [0.308, 0.692, 0.678]$ (Planck Collaboration et al., 2016b) unless explicitly stated.

In this model, the cold dark matter particles respond first to the pull of gravity while baryonic matter follows only after Recombination at $t \sim$400,000 years, when is has had time to cool. The emerging Cosmic Web is therefore originally composed solely of dark matter. The evolution of the dark matter density fields was at first governed by the linear growth of perturbations, a well-studied topic (see e.g., Dodelson 2003). Soon dark matter clumps start collapsing under their own gravity, forming dark matter cores which then grow following the Press-Schechter model (Press & Schechter, 1974; Kauffmann & White, 1993; Lacey & Cole, 1993). Conversely, under-dense regions feel the asymmetric gravitational pull of nearby over-densities and further lose mass, thereby expanding. The intrinsic spheroidal shape of those voids, and their typical size, are governed by the primordial density power spectrum and are thus sensitive to cosmological parameters such as the fractions of baryonic to dark matter density, number of neutrino species and inflationary models. This collapse also gives rise to a filamentary structure of over-densities linking together haloes: the Cosmic 'Web' (Davis et al., 1985; Bond et al., 1996).

Once baryonic matter has cooled sufficiently, it will settle in the potential well of the dark matter which is ~ 6 times more abundant. The gas accumulates within dark matter haloes which become the loci of the first proto-galaxies, while matter in under-dense regions and filaments is left to cool further and will later provide a reservoir of cold gas that will feed galaxies over billions of years (e.g., Kereš et al. 2005; Dekel & Birnboim 2006; Davé et al. 2013).

This simple picture leads to some remarkably robust predictions. Some of our most stringent constraints on the properties of dark matter have come from observations of the clumping of matter in the photo–ionised inter–galactic medium (IGM). In the absence of star formation and associated processes, baryonic matter follows the gravitational potential of the dark matter extremely closely on scales larger than the Jeans length of the gas. Dark matter has also been classed as cold, warm or hot depending on the random velocities of the dark matter particles imprinted in the early Universe. Cold dark matter cannot dissipate energy via radiative processes (Blumenthal et al., 1984). "Warmer" dark matter possesses a larger possible minimal clumping scale due to the "free streaming" of the dark matter particles. Measurements of the free-streaming length of the dark matter particles have been conducted in great detail using the so-called 'Lyman-α forest', the absorption due to neutral hydrogen along a line of sight (e.g., Bode et al. 2001; Macciò et al. 2009; see also Lovell et al. 2011). This has led for instance to constraints of $E_{DM} > 5.3$ keV for a neutrino-like thermal relic dark matter particle (Viel et al., 2013; Iršič et al., 2017).

On larger scales, the clustering of galaxies provides an interesting way to test cosmology. Although they are only imperfect tracers of the underlying dark matter density field, the preferential spacing between objects throughout the Universe can allow us to measure relics of the frozen–out baryonic acoustic oscillations (BAOs; e.g., Tegmark et al. 2006; Alam et al. 2016). Extracting information from BAO surveys is challenging due to the importance of non-linear structure formation effects, such as non-linear halo bias (see Section I.3.1) and redshift space distortions arising from the peculiar velocities of galaxies (Jackson, 1972; Kaiser, 1987). Nev-

ertheless, the cosmological constraints obtained in this way could be the next major step in the post–Planck era.

I.3 Galaxy Formation

The most relevant prediction of the Λ-CDM model is that galaxies should grow hierarchically. The emerging picture is one where galaxies are 'seeded' at the core of dark matter haloes by the pristine gas originating in the surrounding filaments, which has cooled down to form molecular hydrogen (H_2) enabling star formation. As the surrounding gas is accreted, and the host haloes follow cosmic flow, proto–galaxies emerge through a chaotic process of gas-rich mergers (extensively simulated; see e.g., Dubois et al. 2011). In the following phase, galaxies grow by accreting infalling stellar systems and gas (e.g., De Lucia & Blaizot 2007; Khochfar & Silk 2009; Oser et al. 2010; Jones et al. 2017).

Some aspects of the morphologies of galaxies can be attributed to the model of hierarchical growth, although baryonic effects are crucially important as we will see later on. Recent cosmological zoom-in simulations including the effects of stellar age and metallicity predict that the initial collapse of baryons produces mostly 'disk-like' early-type galaxies, but that some primordial, slow rotating, small elliptical galaxies will also form directly from large gas inflows (Kereš et al., 2005; Dekel et al., 2009; Naab et al., 2014). Observations are increasingly supporting this view as they find that roughly $\sim 12\%$ of early-type galaxies are such ellipticals and $\sim 86\%$ are 'disk-like' fast rotators (Krajnović et al., 2011; Emsellem et al., 2011). A 'merger tree' of galaxy formation then unfolds as fast rotators are capable of creating new elliptical galaxies by merging (Heyl et al., 1994; Hoffman et al., 2010; Moster et al., 2011). The eventual merging of the largest elliptical galaxies at late times is a robust prediction of ΛCDM even in the absence of baryonic effects (Kauffmann, 1996; Hopkins et al., 2010).

I.3.1 From the halo–mass function to the stellar–mass function

Dark matter is relatively easy to simulate numerically since gravity is the only force relevant to its dynamics, and baryonic matter only influences it minimally. We are able to predict the distribution and the inner structure of dark matter haloes with great precision (e.g., Aquarius simulations; Bullock et al. 2001). However, feedback processes play a crucial role in shaping the fate of baryons within galaxies as well as putting limits on the cosmological knowledge extractable from galaxy surveys. Unfortunately, measured dark matter halo masses are available for very few galaxies, and modeling their distribution statistically necessitates inferring the halo mass based on the visible matter. The observed distribution of galactic stellar masses (by e.g., Stark et al. 2009; González et al. 2011) is of a very different shape from the predicted mass function of dark matter haloes. Bridging this gap by formulating a reliable halo–mass relation has been a long-standing problem (see two different approaches outlined by Press & Schechter 1974 and Bardeen et al. 1986). The mismatch between the dark matter mass function and observed galactic masses is so large that in the past it has been seen as a threat to the Λ-CDM paradigm (e.g., Boylan-Kolchin et al. 2012). The problem is two-fold: dark-matter only simulations predict too many low-mass haloes compared to the drop-off seen in the galaxy mass distribution at the low mass end (the *missing satellites problem*; Klypin et al. 1999; Moore et al. 1999) and too many haloes are produced by simulations at late times ($z < 2$; the *too big*

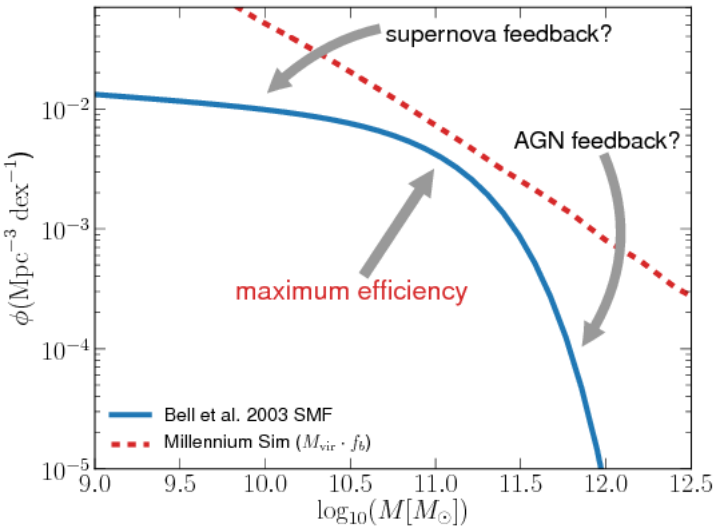

Figure I.2 – Schematic representation of the halo–mass relation. The distribution of dark matter halo masses, multiplied by the cosmic baryon fraction is plotted in dashed red. An analytic fit to observations of the galaxy mass function is plotted in solid blue. Star formation is believed to be limited by stellar feedback at the low halo mass end, and by AGN feedback in the largest haloes. Plot reproduced from Mutch et al. (2013).

to fail problem: Boylan-Kolchin et al. 2011).

In recent years a lot of progress has been made in reconciling these issues by recognising the crucial role played by baryonic feedback in simulations of galaxy populations (e.g., White & Frenk 1991). The modern view of galaxy formation and evolution is driven by the conflict between inflowing, cold gas accretion and outflowing feedback-driven winds (e.g., Shen et al. 2012). Two sources of feedback act at different ends of the mass scale: star formation and active galactic nuclei (AGN). At the high mass end, feedback from a central accreting supermassive black hole (SMBH) has been shown to successfully halt star formation by suppressing gas cooling, perhaps through (radio) jets or X-ray emission (e.g., Croton et al. 2006; Bower et al. 2006; Matzeu et al. 2017). Although the physics of accretion near the centre of the AGN remain debated (e.g., Curtis & Sijacki 2016; Yong et al. 2017), some credence is lent to this idea by the $M - \sigma$ relation: the measured masses of central SMBHs are seen to correlate tightly with the velocity dispersion measured in the bulge of galaxies (Magorrian et al., 1998; Gültekin et al., 2009). Since the bulge is typically $3 - 4$ orders of magnitude more massive than the SMBH, this relation has been interpreted as a sign of co-evolution of galaxies and central black holes. AGN feedback seems necessary in order to prevent the formation of extremely large galaxies at $z < 2$ (e.g., Schaye et al. 2010; Davé et al. 2011).

The suppression of star formation in the smallest galaxies is believed to be caused by stellar winds (Dekel & Silk, 1986). Observations support the view that star formation rate (SFR) correlates with galactic wind speed (Martin, 2005, 2006; Weiner et al., 2009; Chen et al., 2010b; Heckman et al., 2011). This could lead to lower star-forming efficiencies in low-mass galaxies via multiple channels: if the winds are less efficient in small galaxies, star formation could become quenched by low cooling rates in hot, low density gas (see e.g., Schaye et al. 2010). Alternatively, slower gas outflows could mean the winds have a higher mass loading, efficiently ejecting the gas from the galaxy (e.g., Choi & Nagamine 2011; Oppenheimer et al.

2010; Okamoto et al. 2010; but see Bouché et al. 2012). These modes of stellar feedback, over-heating and ejection, have been reasonably successful in reproducing the observed distribution of galaxy masses despite not actually resolving stellar feedback on the physical scales of star clusters. Concerns have been raised about the effects of particular implementations of supernova-driven feedback (e.g., Puchwein & Springel 2013; Smith et al. 2017). Indeed, numerical models choose between injecting the energy produced by supernovae into the galactic environment either in the form of thermal energy or kinetic energy. Due to the lack of resolution of the relevant scales, the former choice effectively counteracts gas cooling, but does not drive outflows, while the latter does the opposite. The division of energy between these two modes is mostly unconstrained because the observational effects are largely degenerate. This has led to attempts to learn about stellar feedback from other probes such as metal enrichment in the circumgalactic medium (CGM; e.g., Keating et al. 2015; Suresh et al. 2015; Bird et al. 2016).

I.3.2 Enrichment of the IGM

Many of the modelling choices used for numerical simulations have revolved around empirically reproducing observations of the stellar mass function at $z \leq 6$. In addition to removing star-forming gas from the haloes, such winds will pollute the environment around galaxies with the metals produced during star formation. Indeed, detections of enriched gas via nebular emission lines show very high speeds and star formation at all redshifts (Rupke et al., 2005; Weiner et al., 2009; Steidel et al., 2010; Coil et al., 2011; Bradshaw et al., 2013; Karman et al., 2014). Metals are observationally detected through the use of intervening absorption lines in quasar absorption spectra (see Section II.4) which provide information about the temperature, kinematics, and ionisation state of the enriched gas – more observables for numerical simulations to match. Different feedback implementations can thus be constrained by comparing the simulations which include the tracking of metals with improved observations of the high redshift CGM metal enrichment.

Numerical models predict different distributions and compositions of gas depending on implementation. The metal enrichment of the CGM and IGM is influenced by a range of factors including the mass distribution of the first generation of stars (Pop III IMF; Pallottini et al. 2014), the activity of the central AGN (e.g., Tescari et al. 2011; Suresh et al. 2015), the ratio of metallicity of the galacitc wind driven by stellar and/or AGN feedback to the metallicity of the host galaxy (Bird et al., 2016; Suresh et al., 2015), and the density threshold for star formation (Finlator et al. 2015; see discussion in Keating et al. 2016). Various observables have been reproduced with some success: for instance the average size of the strongly ionised region around galaxies is of order of a few kiloparsecs (kpc), up to a few hundred kpc (Simcoe et al., 2006; Steidel et al., 2010) and contains large amounts of O VI (Tumlinson et al., 2011) indicating a strong ionising field. Multiple simulations have linked these observations to stellar outflows (Oppenheimer & Davé, 2006; Oppenheimer et al., 2012; Wiersma et al., 2010; Hummels et al., 2013; Barai et al., 2013). Observations of lowly ionised gas using C II and O I have also been reproduced by a range of simulations with moderate success (see review in Keating et al. 2016). These systems are understood to arise from clumps of cold self-shielded gas located close to galaxies (e.g., Finlator et al. 2013). Predictions for metal enrichment further away from galaxies are less reliable. Models disagree on how easy (or hard) it is to expel metals from galaxies fast enough to escape 'gas recycling'. There is disagreement on whether this low-density ex-

pelled gas is cold (Shen et al., 2012, 2014) or hot and highly ionised (e.g., Ford et al. 2014; Suresh et al. 2017), as well as down to which density metal enrichment occurs (Oppenheimer et al., 2009).

Keating et al. (2016) have highlighted the fact that not a single current feedback implementation in numerical simulations satisfactorily reproduces the abundance of the ion C IV[1] at $z > 5.0$. While a few simulations have succeeded at lower redshifts (Oppenheimer & Davé, 2006; Oppenheimer et al., 2012; Rahmati et al., 2016), commonly adopted galactic wind prescriptions fail to enrich the intergalactic medium sufficiently to match the high-redshift measurements incidence rate and/or column density distribution. A similar effect is seen for the ion Mg II. The observed dispersions of these metals seems to require high wind speeds and low mass loading of those winds – properties which are known to fail to reproduce the stellar mass function (Oppenheimer & Davé, 2006; Puchwein & Springel, 2013). Another possibility is that the ionising ultra-violet background (UVB) resulting from the sum total of sources is stronger than expected at $z > 5$, which would allow enriched gas to be more easily detected in C IV. (Finlator et al. 2015 implies the effect could be of order a factor of two; also Oppenheimer et al. 2009). Alternatively, the morphology of early galaxies could be such that wind escape is facilitated through 'funnels'. Testing these hypotheses, and the interplay of galaxy morphology with other factors, would require cosmological simulations with far greater resolution than currently available. At the moment, these processes can only be tested in simulations of individual haloes (e.g., Hopkins et al. 2012, 2014) and the sub-grid models governing the multiple phases of the interstellar medium and feedback are still quite crude.

I.3.3 Direct Observations

While constraining the formation of early galaxies through simulations has been instructive, direct confirmation from observations is invaluable for checking our models. Recent years have seen a push to constrain galactic properties at higher and higher redshifts thanks to the availability of more sensitive instruments such as new instrumentation on the Hubble space telescope (e.g., Jiang et al. 2016a), the IRAC imager on the *Spitzer* space telescope (Bohlin et al., 2011) as well as ground-based instruments including MOSFIRE on the Keck telescope (McLean et al., 2010) and the Atacama Large Millimeter/submillimeter Array (ALMA; Wootten & Thompson 2009). Dozens of galaxies have now been detected at $z > 7$. The brightest of those galaxies are uncovered through the Lyman break technique (Steidel et al., 1996, 2003; Bouwens et al., 2011), relying on their continuum emission to drop sharply at $\lambda \lesssim 1200$Å as absorption by intergalactic hydrogen strongly suppresses it (Lyman-break galaxies or LBGs; see Section II.3.3). The properties of these high-redshift galaxies are then analysed through spectral-energy-distribution (SED) fitting (Caputi et al., 2011; Ilbert et al., 2013) to extract information on the stellar populations, ages, and dust content. However, these techniques cannot be applied to faint sources without a huge investment of resources (e.g., Ouchi et al. 2010; Schenker et al. 2013) or the serendipitous help of gravitational lensing (e.g., Lotz et al. 2017; Hoag et al. 2017). Instead, faint sources are detected through narrow-band imaging, measuring the difference in flux between the object's continuum and redshifted common emission lines. (e.g., Nilsson et al. 2009; Dawson et al. 2007; Trainor et al. 2015; Matthee et al. 2016). This allows a large area of sky to be searched at once, usually for the Lyman-α emission line.

[1] In the notation used throughout this thesis, the transition named C IV occurs from the ion C^{3+}, and so on.

The Lyman break is a necessary feature of any collection of stars surrounded by the IGM. In contrast, it is not entirely clear what causes a galaxy to be a Lyman-α emitter (LAE). The Lyman-α emission line has been found to be more common and stronger in low-mass galaxies (Oyarzún et al. 2016; see also Kornei et al. 2010) and low-luminosity systems (Stark et al., 2010; Forero-Romero et al., 2012), with less dust than the average LBG (Atek et al. 2014; see Laursen et al. 2009). The same trend is seen at lower redshifts (Cowie et al., 2011; Henry et al., 2015), where present-day analogues to LAEs have been used to test the correlations between emission features and stellar properties (e.g., Amorín et al. 2015; Vanzella et al. 2016). The emerging picture is that LAEs are largely young, low-mass galaxies with relatively little dust. If this is indeed the case, those are precisely the kind of objects one would expect to find at very early times according to the hierarchical model: small-size blobs of young, dust-free stars growing into large, more complex galaxies by merging. We would therefore naively expect LAEs to make up an increasingly large fraction of galaxies towards higher redshifts, which does happen, at least until redshift $z \sim 6$ (Ouchi et al., 2008; Cassata et al., 2011; Pentericci et al., 2011; Curtis-Lake et al., 2012; Schenker et al., 2012; Cassata et al., 2015). Beyond $z \sim 6$, the number of LAEs is seen to decline dramatically: only nine solid detections of Lyman-α emission lines have been made at $z > 7$ (Kashikawa et al. 2011; Ono et al. 2012; Pentericci et al. 2014; see Stark et al. 2017 for a listing of current detections). It is believed that this effect is not caused entirely by a drastic change in the stellar populations of early galaxies, but rather by a decrease in visibility of Lyman-α due to ambient cosmic hydrogen that is increasingly neutral: we have entered the epoch of reionisation at this redshift.

I.4 Reionisation

At the conclusion of recombination, almost all electrons in the Universe have been captured by atomic nuclei, forming neutral hydrogen and helium. Some time later, the UV radiation from the first stars started tearing electrons away from these neutral atoms to leave behind an ionised IGM. This process is called reionisation.

I.4.1 Observational Probes

The unfolding of reionisation is closely tied to the nature of the first objects and their ionizing emissivity. Roughly speaking, the first stars and AGN started 'carving' out regions of ionised hydrogen around themselves as soon as they turned on, their influence being at first restricted to their immediate surroundings. This is the classic picture of 'inside–out' or 'bubble–driven' reionisation (Furlanetto & Oh, 2005; Iliev et al., 2006; McQuinn et al., 2007; Trac & Cen, 2007). The visibility of LAEs is strongly suppressed in this regime as the Lyman-α light from the host source does not redshift significantly before hitting the bubble 'wall' and is being resonantly scattered. However, several processes can work against this effect. For instance, line shifts caused by velocity offsets between the object and the moving gas generating the Lyman-α emission can help the light shift out of resonance (Dijkstra et al., 2011; Barnes et al., 2011; Jeeson-Daniel et al., 2012). Clustering of sources can also enlarge the group's bubble faster (e.g., Furlanetto et al. 2006). Nevertheless, the observed decline in LAEs in the redshift range $6 < z < 7$ was strong enough to lead to first estimates the volume filling factor of neutral gas of 40-90% at $z = 7$ to explain their lack of visibility (e.g., Ciardi et al. 2012; Jensen et al.

2013). However, a fast change in the escape fraction of ionising radiation from the hosts – a quick increase in how porous they are – could explain the fast evolution of LAE visibility (e.g., Mesinger et al. 2015; Choudhury et al. 2015). Other contributing effects which have been suggested include an increase in LAEs' stellar masses and star formation rates leading to a fast increase in UV emission (e.g., Bouwens et al. 2007); the rapid destruction and ionisation of dust to enable higher escape fractions (e.g., Forero-Romero et al. 2012); or it is possible that LAEs are being hidden by optically thick absorbers at the end of reionisation (Bolton & Haehnelt, 2013).

One significant consequence of reionisation is that the photons originating from the CMB will scatter off of free electrons once the Universe is ionised. This is measured by the Thomson optical depth to reionization, $\tau_e = \sigma_T \int_{\text{present}}^{\text{CMB}} dl\, n_e(r)$ where n_e is a number density of free electrons along the line of sight, σ_T is the Thomson cross section and l is the comoving distance. τ_e quantifies the photon mean free path to reionisation, which is proportional to the average number of free electrons that a CMB photon has undergone Thompson scattering with. This process diminishes the amplitude of the CMB $\langle \text{TT} \rangle$ power spectrum by a factor $\propto \exp(-2\tau_e)$. Polarisation of the CMB also arises due to the differential scattering by the electron gas of photons originating from hot and cold patches; this creates characteristic features in the CMB $\langle \text{TE} \rangle$ and $\langle \text{EE} \rangle$ polarisation cross–spectra (see Zaldarriaga 1997). The Planck satellite measured $\tau_{\text{e}} = 0.058 \pm 0.012$ (Planck Collaboration et al., 2016b). In a simplified model where reionisation occurs instantaneously, this optical depth corresponds to this happening ~ 500 My after the Big Bang or at $z = 8.8^{+1.0}_{-1.1}$. Note that a measurement of τ_{Re} tells us nothing about the start or the unfolding of reionisation. Also note that the value of τ_{Re} has declined significantly since the 2011 measurement by the WMAP satellite ($\tau_{\text{Re}} = 0.088 \pm 0.014$, Hinshaw et al. 2013), explaining why older literature was striving to reconcile other observations with a very early ($z \sim 15$) start of reionisation (Bromm & Yoshida, 2011; Dunlop, 2013).

The present-day Universe is highly ionised with abundances of neutral hydrogen in the IGM of $\Omega_{\text{HI}} = 4 \cdot 10^{-4}$ corresponding to hydrogen neutral fractions $f_{\text{HI}} \sim 10^{-5}$ (Zwaan et al., 2005; Martin et al., 2010; Delhaize et al., 2013). The rare occurrences of neutral gas are mostly in the form of dense clouds with little or no star formation and metal enrichment, which are believed to be cold inflows originating in the voids and filaments of the cosmic web (e.g., Péroux et al. 2005; Davé et al. 2013). Because these systems are faint, they were first seen as absorption systems in the Lyman-α emission of distant quasars – earning them the name of Damped Lyman-α systems (DLAs). The amount of neutral hydrogen represented by DLAs declines by a factor of ~ 2 between $z = 5$ and the present day, evolving as $(1 + z)^{0.4}$ (Noterdaeme et al. 2012; Crighton et al. 2015; but see Songaila & Cowie 2010). This is consistent with the picture of these last 'pockets' of neutral gas being converted into stars or being destroyed by the UVB as time goes by. DLAs also increase in metallicity with decreasing redshift, indicating a fast change in either their metal enrichment processes or morphology (Wolfe et al., 2005; Rafelski et al., 2014). Beyond $z = 5$, the Lyman-α forest saturates and does not allow for DLAs to be detected using conventional techniques (see Section II.3.3 for more details).

Finally, the most precise measurements of the neutral fraction of hydrogen during and just after reionisation have come from analysing the Lyman-α forest of quasars (Becker et al., 2001; Fan et al., 2002; White et al., 2003; Songaila, 2004; Fan et al., 2006). By measuring the mean Lyman-α opacity (see Chapter 2), Fan et al. (2006) showed that the IGM reaches a global neutral fraction of $\sim 0.1\%$ around $z \sim 6$, increasing from low redshift. Becker et al. (2015b) extended this study to include 42 quasars at $z > 4.5$, confirming the accelerating evolution in

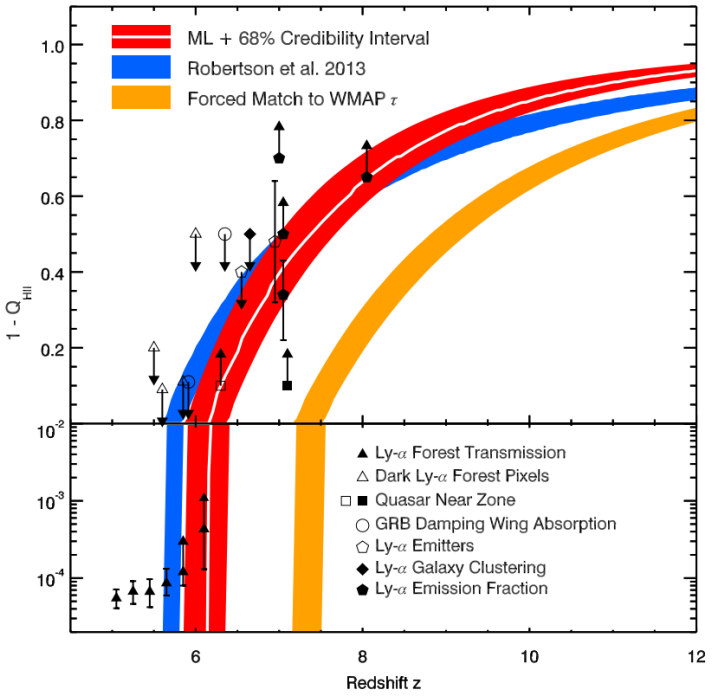

Figure I.3 – Summary of current constraints on the hydrogen neutral fraction. In addition to the results mentioned in the text, the gamma ray burst (GRB) measurements from Chornock et al. (2013) are included. The red and blue lines are models computed by Robertson et al. (2015) and Robertson et al. (2013), respectively. Reproduced from Robertson et al. (2015).

the redshift range $5.5 < z < 6.0$ and discovering a $\sim$ 110Mpc h^{-1} Lyman-α trough extending down to $z = 5.5$ towards the $z = 6.0$ quasar J0148+0600. The source of the volume-averaged opacity evolution is generally understood to be a weakening of the UVB (e.g., Bolton & Haehnelt, 2007b, Calverley et al., 2011, Wyithe & Bolton, 2011). However, the origin of the opacity variations between sightlines is the object of much debate and has given rise to three families of models which attempt to match both the mean evolution of the neutral fraction and its scatter. Chardin et al. (2015) proposed that a large scatter arises naturally if bright rare sources (such as AGN) are a significant contributor to reionisation (see also Chardin et al. 2017b). This ties in with concerns about galaxies potentially not being sufficient to reionise the Universe as mentioned earlier. Secondly, D'Aloisio et al. (2015) argued for the inclusion of temperature effects into numerical simulations. In their model, temperature fluctuations arise from patchiness in the timing of reionization. High-density regions reionise first and have more time to cool, producing a higher recombination rate and larger opacity. Low-density regions, in contrast, reionise late and are still hot at $z \sim 6$. Thirdly, Davies & Furlanetto (2016) point out that such scatter in Lyman-α opacity is expected from fluctuations in the UV background due to a short and spatially varying mean free path of ionising photons. Indeed, a slight spread in emissivity can have a strong impact on the UVB due to coupling between the emissivity and the local mean-free path (e.g., McQuinn et al., 2011). The implications of these models and differences between them are an active subject of research.

I.4.2 Galaxies or quasars?

After the end of reionisation, the IGM is kept ionised by the UVB consisting of the total sum of UV emission from stars, AGN, and less important sources such as cosmic rays and hard X-ray binaries. The relative contributions to the UVB from stars and AGN, and their evolution across cosmic time, are a topic of ongoing debate. It is still very much unclear whether galaxies and quasars are responsible for reionisation to comparable extents, or whether either of these two sources are negligible compared to the other; all three stances have been argued in the recent literature.

Traditionally, star-forming galaxies were suspected to be the major or sole contributors to the reionisation photon budget (e.g., Madau et al. 1999; Meiksin 2005; Faucher-Giguère et al. 2009). Recently, however, doubts regarding their abundance in the early Universe and the ability of ionising photons to escape from these sources have made the picture more complicated. The first of these issues has to do with the UV luminosity function (LF) of galaxies at $z > 5$. For typical assumptions about their ionising spectra and the fraction of ionising photons that escape ($f_{\rm esc}$) there are not enough galaxies detected with magnitudes $M_{\rm UV} < -15$ to reionise the Universe (Bouwens et al., 2012; Finkelstein et al., 2012). Extrapolating the luminosity function as a power law below the detection thresholds, however, solves this problem (Haardt & Madau, 2012). Some numerical models of galaxy formation, and some observations, have seemingly shown there to be abundant low-luminosity galaxies (sims: Liu et al. 2016; Gnedin et al. 2016; Ocvirk et al. 2016; Finlator et al. 2017; observations: Livermore et al. 2017), but other models have predicted a turnover luminosity below which objects do not contribute significantly (Jaacks et al., 2013; Yue et al., 2016). Some observations have reported a turnover, leading to confusion and highlighting the difficulty of such measurements (e.g., Castellano et al. 2016). Early galaxies have either been counted through deep field exposures or gravitational lensing. The latter technique makes it possible to probe to very faint intrinsic luminosities but comes with a variety of systematic difficulties. Results from deep fields and gravitational lensing have therefore sometimes not agreed. A review by Bouwens et al. (2016) has shown that when accounting for systematic errors rigorously, the location of the LF turnover and the faint-end power law slope are largely unconstrained at $M_{\rm UV} < -15$.

At the same time, it is unclear how efficiently Lyman continuum radiation (Ly-C) escapes from galaxies. The search for Ly-C at $z \gtrsim 3$ in LAEs has been extremely unsuccessful, with only a few confirmed detections and nearly all of them contested (Mostardi et al., 2015; Siana et al., 2015; Vanzella et al., 2016). There are a few reasons why the search for Ly-C is complicated at high redshift: neutral intergalactic absorbers can attenuate Ly-C (Inoue et al., 2014), lower redshift interlopers are common (Vanzella et al., 2010), the geometry of early galaxies evolves quickly, adding stochasticity (Wise et al., 2014; Cen & Kimm, 2015), and selection effects favour the brightest sources which are suspected to have lower escape fractions than lower-luminosity galaxies (Grazian et al., 2016). To counter these problems, Ly-C leakage has been studied in local Universe analogues instead with more success (Borthakur et al., 2014; Izotov et al., 2016). It is not clear how well these sources correspond to reionization-era galaxies or whether the values of $f_{\rm esc}$ measured locally are even high enough for galaxies to reionise the Universe alone (Vanzella et al., 2012; de Barros et al., 2016; Mas-Ribas et al., 2017). Hope might come from observations showing that high redshift galaxies have stronger O III emission than low-redshift ones (Zitrin et al., 2015). The rare $z \sim 3$ confirmed Ly-C leakers all have this feature (e.g., Finkelstein et al. 2013; Roberts-Borsani et al. 2016). O III emission is predicted

by models to correlate with starbursts and therefore with a higher $f_{\rm esc}$ (Jaskot & Oey, 2013; Nakajima & Ouchi, 2014). This has also been confirmed locally (James et al., 2016). With reasonably large (optimistic) values of both the UV LF faint end parameters and $f_{\rm esc}$, galaxies are capable of providing all reionisation photons without help from quasars.

Interestingly, it appears that within observational constraints early quasars may be able to drive reionisation without help from galaxies as well (see Madau & Haardt 2015). Ionising radiation from the accretion processes happening in AGN are believed to easily carve out their host galaxies' neutral gas and dust, perhaps through jets, so that virtually all of the radiation can escape. The deciding factor for the contribution of quasars to the reionisation budget is therefore their (ionising) luminosity function. Unfortunately, again, the different methods employed to identify faint AGN are diverse, affected by complex systematics, and do not agree with each other (Madau & Haardt, 2015; Khaire & Srianand, 2015). Only recently have some measurements been consistent with a quasars-only model (Glikman et al., 2011; Fiore et al., 2012; Giallongo et al., 2015), and their reliability is debated (Kim et al., 2015; Qin et al., 2017; Parsa et al., 2017).

Whether galaxies or AGN are driving reionisation has far-ranging consequences. One important aspect ties in to He II reionisation and the evolution of IGM temperature. When an atom becomes ionised, any excess energy is turned into thermal excitation, resulting in heating (e.g., Hui & Gnedin 1997). He II in the IGM has a higher ionisation potential than hydrogen and is seen to undergo reionisation at $z \sim 2.5$, long after hydrogen (Miralda-Escudé, 2000). Traditionally, the high-energy photons produced by quasars have been held responsible for ionising He II for two reasons. Firstly, the number density of quasars and their X-ray emission is observed to peak at $z \sim 2.5$, just in time to trigger He II reionisation (Ueda et al., 2003; Furlanetto & Oh, 2008; Faucher-Giguère et al., 2009; Haardt & Madau, 2012). Secondly, the temperature of the IGM is measured to increase dramatically over the He II reionisation era, consistent with an increase in hard UV photons which are not thought to come from galaxies (e.g., McQuinn et al. 2009; Shull et al. 2010; Compostella et al. 2013). This seems to imply that quasars could not have ionised hydrogen at $z > 6$ since He II would have been reionised simultaneously and the IGM would already be heated by $z \sim 2.5$. However, recently He II lines-of-sight displaying low He II opacity have been found as early at $z \sim 3.5$ (Worseck et al., 2016), indicating that the processes responsible might have been at work earlier than thought. Uncertainties in measurements of the IGM temperature are still large, not excluding an extended He II reionisation period starting as early as $z \sim 6$ (Bolton et al., 2012; Boera et al., 2014; Madau & Haardt, 2015) and alternative models for He II reionisation and IGM heating not caused by quasars have been proposed (e.g., Puchwein et al. 2012). Taken together, the evidence does not rule out an action of quasars on hydrogen reionisation earlier than previously thought.

In the next Chapter, we go into more detail about the ways quasars provide information on the early Universe. Since they are the brightest persistent sources of radiation in the Universe, the properties of quasars and their evolution across cosmic time have been studied in great detail. Quasars are powerful probes of reionisation as well as potentially agents driving the process.

II

Quasar spectroscopy to uncover Cosmic Dawn

You can't do astrophysics just by taking pictures through little colored pieces of glass.

Philip Massey

Abstract

THIS chapter reviews technical details relating to the use of background quasars as probes of cosmic structure. Extracting information from these objects requires spectra taken with large aperture telescopes and efficient data reduction. The most widely used atomic transitions in the spectra of high-redshift quasars are the Lyman series, including the Lyman-α forest. Other transitions originating from metals along the lines of sight to quasars provide constraints on metal enrichment and the state of the UVB across redshift.

II.1 Introduction

Across cosmic time, accretion onto a galaxy's central supermassive black hole can produce radiation that is orders of magnitude brighter than the combined output from all the stars in the galaxy. This was already the case at $z = 7$, where the quasar ULAS J1120+0641, the most distant currently known, displays an intrinsic luminosity of $L = 6.3 \cdot 10^{13} L_{\odot}$ powered by a $M = 2 \cdot 10^{9} M_{\odot}$ black hole. This enormous brightness means that both the quasar's resonant emission lines and its spectral continuum can be detected even only 700 My after the Big Bang. The early history of the Universe can thus be probed through *absorption spectroscopy*.

If a quasar at redshift $z = z_{\rm em}$ is emitting continuum radiation, by the time the light encounters an intervening structure at redshift $z_{\rm abs}$ along the line of sight it has redshifted with respect to the quasar such that absorbed wavelength, $\lambda_{\rm abs}$, is related to the emitted wavelength, $\lambda_{\rm em}$, as:

$$\lambda_{\rm abs} = \lambda_{\rm em} \frac{z_{\rm abs} + 1}{z_{\rm em} + 1}. \tag{II.1}$$

Thus each time that gas capable of scattering at a particular wavelength is encountered, a corresponding feature is imprinted on the quasar continuum. The availability of bright background sources combined with the expansion of the Universe thus make absorption lines invaluable tools for cosmology.

This Chapter will focus on absorption features in quasar spectra detected at visible and infrared wavelengths from Earth, but it is worth noting the same techniques have been applied at high redshift using gamma ray bursts (GRBs) as background sources. Although rarer and short-lived, valuable information has also been gained from their study (see Tanvir et al. 2009; Totani 2013 and references therein). Absorption lines are also an important tool for the study of He II Reionisation at $z \sim 2.5$, again with quasars as background sources, but the using He II Lyman series instead (see Syphers et al. 2009 and references therein).

First, I will describe some of the technical details related to the data used in this work, and describe how it was obtained and analysed. Next, I will turn to the physics of the hydrogen Lyman series and the many ways it has been used to study the Epoch of Reionisation. Finally, I will turn to other transitions corresponding to heavier elements and metal enrichment.

II.2 Echelle and echellette spectroscopy

Most of the data used in this thesis was obtained with *echelle spectrographs*. In its simplest form, an astronomical spectrograph disperses light coming from an object into a spectrum using a *disperser*, which for the spectrographs used here is a diffraction grating. The light is first isolated through a *slit* and collected by a *collimator*. After the spectral splitting, the dispersed light is focussed using a *lens* onto a *detector*, usually a CCD plate. Compared to a classical spectrograph, an echelle adds a second disperser to scatter the light into separate diffraction orders. The grating equation tells us that,

$$m\lambda = 2\sigma \sin\theta, \tag{II.2}$$

where λ is the wavelength of incoming light, θ is the angle the light will be deflected by, σ is the grating groove density, and m is the order. It can readily be seen that if the incoming light spans a wide range of wavelengths, higher orders will overlap on the detector and information will be lost. A single-order spectrograph gets around this by using *blocking filters* to limit the transmission to only the lowest wavelengths with $m = 1$ or 2. In this work, spectrographs such as GMOS S, LBT MODS, EFOSC2, FOCAS and the MMT RCS are such 'single-channel' spectrographs. In contrast, an echelle disperses several orders (from a few up to several tens), allowing for a much wider wavelength coverage in a single observation. Echelle spectrographs such as HIRES and UVES use a grating for order dispersion. echellette spectrographs on the other hand use a prism, resulting in a characteristic 'curving' of the orders on the detector. X-Shooter, ESI, and MagE are echellette spectrographs.

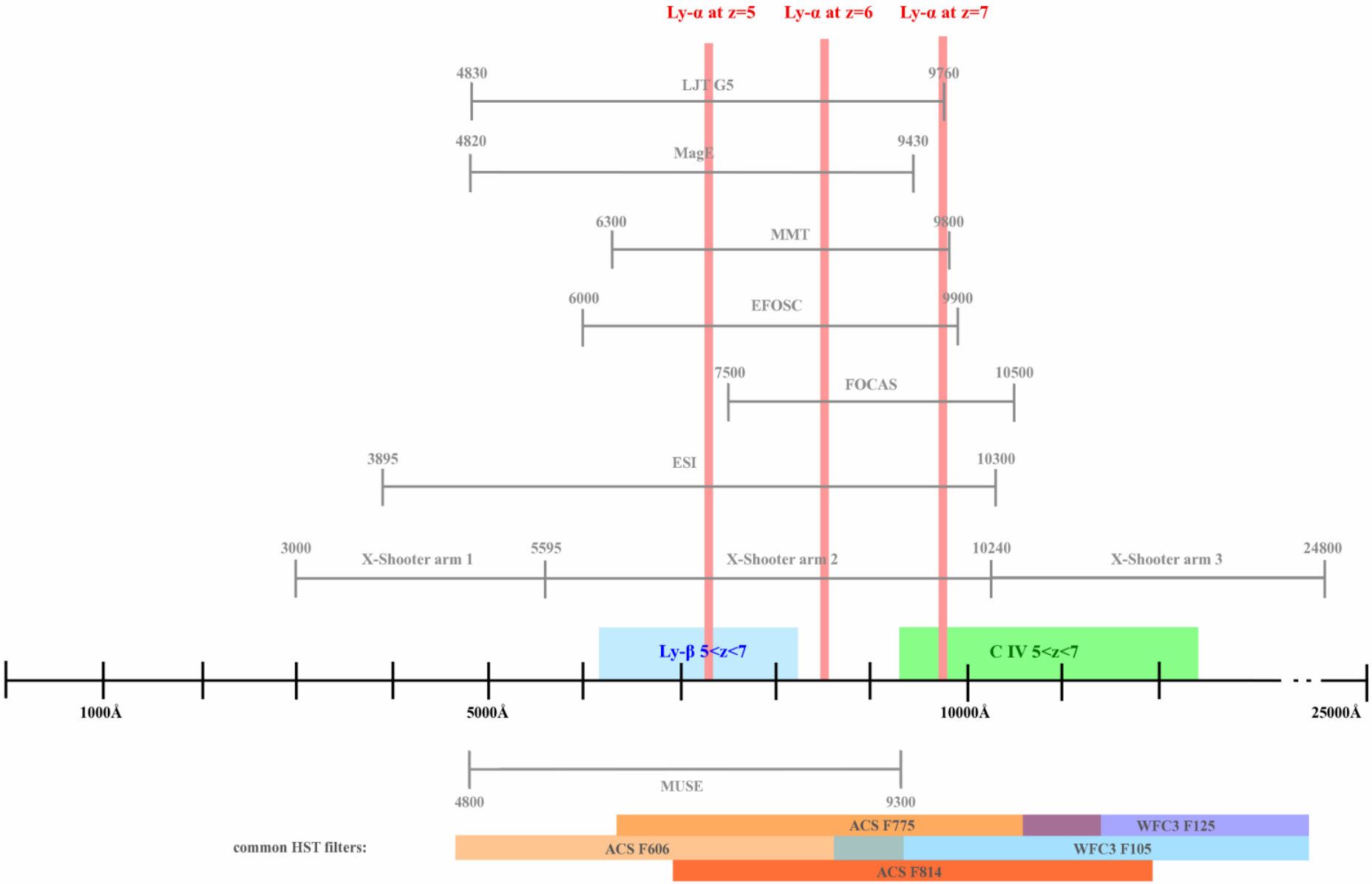

Figure II.1 – Ranges of wavelengths covered by the spectrographs mentioned in the thesis. Wavelengths of physical relevance are also shown, with the range of the MUSE IFU and commonly used HST filters at the bottom. Sensitivity ranges were obtained from the instruments' online manuals; in practice they may vary slightly.

As a trade-off, the slit of an echelle or echellette is shorter than a single-order spectrograph's to avoid transversal overlapping of orders. This complicates the process of sky subtraction as less blank field is available. The introduction of a second disperser compounds any defects in the gratings or prisms, such as anamorphic distortions. For these reasons, reduction of echellette data is generally more complicated than reducing single-order data.

The resolving power of a disperser is characterised by its *resolution* $R = \lambda/\Delta\lambda$, which is wavelength-dependent. For the spectrographs mentioned in this work, the resolution spans the range $4000 < R \lesssim 30000$ around a central wavelength of $\lambda = 7000$ Å, corresponding to an ability to distinguish lines 0.2 Å$\lesssim \lambda < 1.75$ Å apart. The typical throughput of the instruments, including the telescope, is $\sim 20\%$, and the typical grating used has $\sigma =\sim 80$ grooves/mm. For a review of instrumental design details of spectrographs used in astronomy, see Massey & Hanson (2013).

To be more precise, the work presented in Chapter V uses optical spectroscopy obtained on nine instruments. The Gemini Multi-Object Spectrograph (GMOS-S, Hook et al. 2004) is mounted on the 8.1m Gemini South telescope on the summit of Cerro Pachon in Chile. The Multi-Object Double Spectrographs for the Large Binocular Telescope (LBT MODS, Pogge et al. 2012) were designed and built for the twin 8.4-meter diameter mirror Large Binocular Telescope on Mt. Graham in southeastern Arizona. EFOSC-2 – the ESO Faint Object Spectrograph and Camera (Buzzoni et al., 1984) is currently located on the 3.58m New Technology Telescope (NTT) in La Silla which uses adaptive optics. Some spectra originate from the Red Channel Spectrograph on the 6.5 Multiple Mirror Telescope (MMT RCS, Schmidt et al. 1989) in Arizona. The last single-order instrument used here is the Faint Object Camera and Spectrograph on the 8.2m Subaru Telescope (FOCAS, Kashikawa et al. 2002) on Mauna Kea, Hawaii.

The only echelle spectrograph used in this thesis is the High Resolution Echelle Spectrometer (HIRES, Vogt et al. 1994) on the Keck 10m Telescope in Hawaii, which has the highest resolution out of the instruments used here. Echellette spectrographs include the Echellette Spectrograph and Imager, also on the Keck II Telescope (ESI, Sheinis et al. 2002) and the Magellan Echellette (MagE, Marshall et al. 2008) on the Clay 6.5m Telescope in Chile. Finally, X-Shooter on the 8.2m Very Large Telescope (VLT) at the Paranal Observatory (Vernet et al., 2011) is an echellette instrument which covers the optical and near-infrared in three separate arms.

Spectra originating from three more instruments were gathered but not used here: these are the Ultraviolet and Visual Echelle Spectrograph (UVES, Dekker et al. 2000) on the VLT, the Double Spectrograph on the Hale 5.1m telescope in Palomar (DBSP, Oke & Gunn 1982) and the Yunnan Faint Object Spectrograph and Camera (YFOCS) mounted on the Li-Jiang 2.4m telescope.

II.2.1 Data reduction

Chapter 5 of this thesis makes use of new data obtained on the ESI spectrograph as well as new reductions of archival data from various sources, and the extraction procedure is therefore briefly discussed here. The process of extracting usable data from spectrograph detectors consists of five main steps:

1. Exposures of calibration lamps, either internal to the instrument or illuminating the inside of the telescope dome, are used for flat-fielding and to derive a wavelength solution.

2. The spectrograph obtains the spectrum of a previously known standard star in the vicinity of the target object. This establishes the sensitivity across the detector and enables a relative flux calibration.

3. After the observation, sky subtraction is performed. This is done by repeated rejection of bad pixels corresponding to cosmic rays and other defects, and the generation of a best-fit sky model through least squares fitting. For the data reduced in this thesis, the spectrum is fit with a bivariate B-spline that is a function of the wavelengths obtained from step 1.

4. A one-dimensional spectrum is extracted, and spectra from multiple exposures are combined if necessary.

5. If possible, the absolute flux of the resulting spectrum is pegged to pre-existing measurements from other instruments or photometry. This accounts for the unknown value of the total telescope and instrument throughput and/or slit losses.

Additional details on the specific data analysis pipelines and sky subtraction software used for this thesis can be found in the relevant sections (see also Kelson 2003).

II.3 Lyman-α emission and absorption

II.3.1 Background

The *Lyman series* is the name given to a set of atomic transitions of hydrogen located in the rest-frame ultraviolet, discovered in 1906 by Theodore Lyman (Lyman, 1906). The wavelengths corresponding to the Lyman series can be calculated from the Rydberg formula as:

$$\lambda = \frac{hc}{E_i - E_f} = \frac{12398.4 \text{ Å eV}}{E_i - E_f}, \tag{II.3}$$

where E_i and E_f correspond to the initial and final energy levels of the electron respectively. The energy of the n^{th} energy level of the hydrogen atom is given by $E_n = -13.6 \text{ eV}/n^2$ such that,

$$\frac{1}{\lambda} = (911.8 \text{ Å})^{-1} \left(\frac{1}{m^2} - \frac{1}{n^2} \right). \tag{II.4}$$

The Lyman series results from using $m = 1$ in this equation, meaning it arises from electrons dropping to the ground state from a higher energy level (the Balmer series is obtained by using $m = 2$). The longest wavelength, $\lambda = 1215.67$Å, corresponds to $n = 2$ and is denoted with the letter α; further transitions are denoted with the successive greek letters. There are therefore infinitely many Lyman transition lines, although they become closely packed as the *Lyman limit* ($n \to \infty$, $\lambda = 911.8$Å) is approached.

Broadening effects come into play which give the emission (or absorption) a wider spectral distribution. Generally speaking, broadening can be classified into two types: Lorentz-profile broadening arises from the uncertainties in the upper and lower energy levels due to their finite lifetimes. These lifetimes are affected by both radiative and collisional effects, such as collisions with neutral ions inside the gas and close encounters with the electric field of neighbouring ions. In low density environments such as the IGM (and sometimes the interstellar

medium) the primary process which determines energy level lifetimes is spontaneous decay. Lorentz broadening takes the form,

$$\Phi_\lambda = \frac{1}{\pi}\frac{\mu_k}{\mu_k^2 + (\lambda + \lambda_0)^2}, \tag{II.5}$$

where λ_0 is the nominal (central) wavelength of the transition, and the level depopulation effects have been bundled up into the parameter μ_k, sometimes called the radiation damping constant.

The other class of broadening results from the intrinsic distribution of peculiar velocities inside the gas, and is referred to as Doppler broadening. It takes the form of a Gaussian distribution,

$$\Psi_\nu = \frac{1}{\sqrt{\pi}b} e^{-\left(\frac{\nu-\nu_0}{b}\right)^2}, \tag{II.6}$$

where $\nu_0 = c/\lambda_0$ is the central frequency and the velocity effects are contained in the Doppler parameter b. The most well-known example of Doppler broadening is due to the thermal motions of the gas that follow a Maxwell distribution; the resulting Doppler parameter is given by $b_T = \sqrt{2kT/m}$. It is important to note that turbulence effects, both macroscopic and microscopic, can boost the total b as $b = \sqrt{b_T^2 + b_{\text{turb}}^2}$.

The combination of these two effects results in a spectral shape called the *Voigt profile*, which is given by the convolution of these two shape functions, $\phi(\lambda) = \Phi(\lambda) \otimes \Psi(\lambda)$. When the column density is low, the profile is dominated by a Gaussian core of Doppler broadening and a simple Gaussian fit will give a good approximation. Quantum mechanical effects become important at higher column densities, and are responsible for extended profiles with 'damping wings' further from the central wavelength. The Voigt profile does not have an analytic form. For a more thorough treatment, see Pettini (2010); Howarth (2010).

Scattering of incoming light results in a decrease in flux intensity,

$$I_\lambda = I_{\lambda,0} e^{-\tau_\lambda}. \tag{II.7}$$

The optical depth τ is related to the absorption profile as $\tau(\lambda) = N\sigma_0\phi(\lambda)$, where N is the column density of the gas and σ_0 is the specific cross-section of the transition. The *equivalent width*, W, of an absorption line is defined as the width of a unit-height rectangle with the same amount of absorption. This can be calculated as,

$$W = \int_{-\infty}^{\infty} d\lambda \frac{I_{\lambda,0} - I_\lambda}{I_{\lambda,0}} = \int_{-\infty}^{\infty} d\lambda\, 1 - e^{-\tau(\lambda)}. \tag{II.8}$$

We can now see that if τ is small, i.e. in the limit of weak absorption, we can make the approximation

$$W = \int_{-\infty}^{\infty} d\lambda\, 1 - e^{-\tau(\lambda)} \cong \int_{-\infty}^{\infty} d\lambda 1 - 1 + \tau = \int_{-\infty}^{\infty} d\tau\tau \tag{II.9}$$

$$= N\sigma_0 \int_{-\infty}^{\infty} d\tau\phi = N\sigma_0. \tag{II.10}$$

Such lines are said to be in the linear regime of the curve of growth (the relationship between W and N), because W is proportional to N. This is exploited in the *apparent optical depth method*, in which the column density for weak lines is approximated from the optical depth

alone without the need to fit a full Voigt profile. Using $\tau = -\ln(I/I_0)$, the column density is given by

$$N = \frac{W}{\sigma_0} = \frac{m_e c}{\pi e^2} \frac{\tau}{f \lambda_0} \tag{II.11}$$

where we have parametrised σ_0 in terms of the more commonly used *oscillator strength* f (see Savage & Sembach 1991 for more details).

II.3.2 Gunn-Peterson absorption

Gunn-Peterson absorption (Gunn & Peterson, 1965) is the name given to absorption of light from background objects by uniform neutral hydrogen in the intergalactic medium. In the following section, I derive the simple analytic form of this phenomenon by roughly following Gunn & Peterson (1965).

Let a distant quasar be located at a redshift $z = z_0$. The contribution to the optical depth from a slab of intervening gas of proper thickness dl located at z can be written as,

$$d\tau = n(z)\, \sigma(\nu)\, dl, \tag{II.12}$$

where $n(z)$ is the number density at redshift z and $\sigma(\nu)$ is the absorption cross-section, which varies with the frequency of the light ν. To a good approximation (except at very high temperatures), $\sigma(\nu)$ can be expressed as a delta function centred on the resonant frequency ν_α:

$$\sigma(\nu) = \frac{\pi e^2}{mc}\, f\, \delta(\nu - \nu_\alpha) \equiv \sigma_\alpha\, \nu_\alpha\, \delta(\nu - \nu_\alpha). \tag{II.13}$$

Here the oscillator strength of the transition, f, has been folded into the characteristic cross-section σ_α. By the time the quasar's light reaches the slab of gas, its frequency has redshifted to $\nu\,(1+z)$, making the total optical depth for an observer located at $z = 0$,

$$\begin{aligned} \tau = \int_0^{z_0} d\tau &= \int_0^{z_0} n(z)\, \sigma\left(\nu\,(1+z)\right) \frac{dl}{dz}\, dz \\ &= \int_0^{z_0} \frac{n(z)\, \sigma_\alpha\, \nu_\alpha\, c}{H(z)\,(1+z)} \delta\left(\nu\,(1+z) - \nu_\alpha\right)\, dz, \end{aligned} \tag{II.14}$$

where we have used the fact that $dl/dz = cH^{-1}(z)(1+z)^{-1}$.

The Hubble parameter in the matter-dominated era can be approximated as $H(z) = H_0\sqrt{\Omega_m}\,(1+z)^{3/2}$. To solve the above integral, we make use of the delta function by changing the integration variable to $d(z+1)$ and use the identity $\delta(\alpha x) = \delta(x)/|\alpha|$,

$$\begin{aligned} \tau &= \int_1^{z_0+1} \frac{n(z)\, \sigma_\alpha\, \nu_\alpha\, c}{H_0\, \sqrt{\Omega_m}\, (1+z)^{5/2}\, \nu}\, \delta\left((1+z) - \frac{\nu_\alpha}{\nu}\right) d(z+1) \\ &= \left.\frac{n(z)\, \sigma_\alpha\, \nu_\alpha\, c}{H_0\, \sqrt{\Omega_m}\, (1+z)^{5/2}\, \nu}\right|_{\nu\, =\, \nu_\alpha/(1+z)} \\ &= \frac{n(z)\, \sigma_\alpha\, c}{H_0\, \sqrt{\Omega_m}\, (1+z)^{3/2}}. \end{aligned} \tag{II.15}$$

In the context of reionisation, it is convenient to express the mean number density of neutral hydrogen, $n_{H\ I}(z)$, as a fraction of the mean total hydrogen number density $\bar{n}_H(z)$,

$$n_{H\ I} = \langle x_{H\ I} \rangle \, \bar{n}_H. \tag{II.16}$$

$\bar{n}_H(z)$ can be calculated from the cosmological critical density[1],

$$\bar{n}_H = \rho_{\text{crit}} \frac{X}{m_p} \Omega_b (1+z)^3 \qquad \rho_{\text{crit}} = \frac{3H^2(z)}{8\pi G} \tag{II.17}$$

where $X = 0.76$ is the primordial hydrogen fraction, $m_p = 2.39 \cdot 10^{-18}$ kg is the proton mass, and the other constants have their usual meaning. Combining those yields $\rho_{\text{crit}} = 9.06 \cdot 10^{-27}$ kg m^{-3}, and with the resonant cross-section of the Lyman-α transition $\sigma_\alpha = 4.48 \cdot 10^{-22}$m^2, we finally get:

$$\tau = 3.2 \cdot 10^5 \, \langle x_{H\ I} \rangle \qquad \text{at } z = 5 \tag{II.18}$$

or in a more flexible form:

$$\tau = 3.2 \cdot 10^5 \, \langle x_{H\ I} \rangle \left(\frac{X}{0.76} \right) \left(\frac{\Omega_m h^2}{0.142} \right)^{-0.5} \left(\frac{\Omega_b h^2}{0.022} \right) \left(\frac{1+z}{6} \right)^{1.5} \tag{II.19}$$

Since the transmitted flux scales exponentially with optical depth as $F = e^{-\tau}$, even a tiny fraction of 1 neutral atom in 50,000 is sufficient to fully absorb the flux coming from the background quasar. Observationally, these 'Gunn-Peterson troughs' occur at redshifts $z = 5.5$ and beyond (Fan et al., 2001; White et al., 2003). The presence of transmitted flux in quasar spectra is a strong indication that reionisation had concluded – in a globally averaged sense – by the end of the first billion years of the Universe.

II.3.3 Lyman-α forest and Lyman-α transmission

At redshifts lower than $z \sim 5.7$ where Lyman-α absorption by the IGM is not saturated, individual Lyman-α absorption features occur whenever a quasar line of sight passes through a clump of gas. The accumulation of these features across a large redshift range gives rise to the *Lyman-α forest*, a powerful tracer of neutral hydrogen (Lynds & Stockton, 1966; Burbidge et al., 1966; Weymann et al., 1981). If the column density of the intervening gas cloud exceeds $N(HI) \gtrsim 10^{17}$cm^{-2}, a discontinuity in flux will be observed at the Lyman limit in the rest-frame of the absorber due to the overlap of high-order Lyman lines. If $N(HI) \gtrsim 10^{19}$cm^{-2}, the absorption feature will display prominent wings – the origin of the naming of 'Damped' Lyman-α systems[2] (Wolfe et al., 1986; Rauch, 1998; Miralda-Escudé, 1998).

Some of the physical properties of the absorbing gas clouds are well understood. For instance, their typical temperature is $T \sim 10^4$K, and they are photo-heated and photo-ionised by the UVB (Croft, 2004). The absorbing gas is also metal-poor and may be pristine at very low densities, although it is unclear whether the few metals it contains are produced by star formation within the clouds or brought by IGM stirring (e.g., Theuns et al. 2002; Schaye et al. 2003; McDonald et al. 2005; Simcoe et al. 2006; Rafelski et al. 2014). The distribution of low column

[1] This analysis can be generalised to the case where the density field is non-uniform. This is known as the 'fluctuating' Gunn-Peterson approximation.

[2] In the literature, systems with $10^{19} \lesssim N(HI) \lesssim 10^{20.3}$cm^{-2} are sometimes referred to as sub-DLAs.

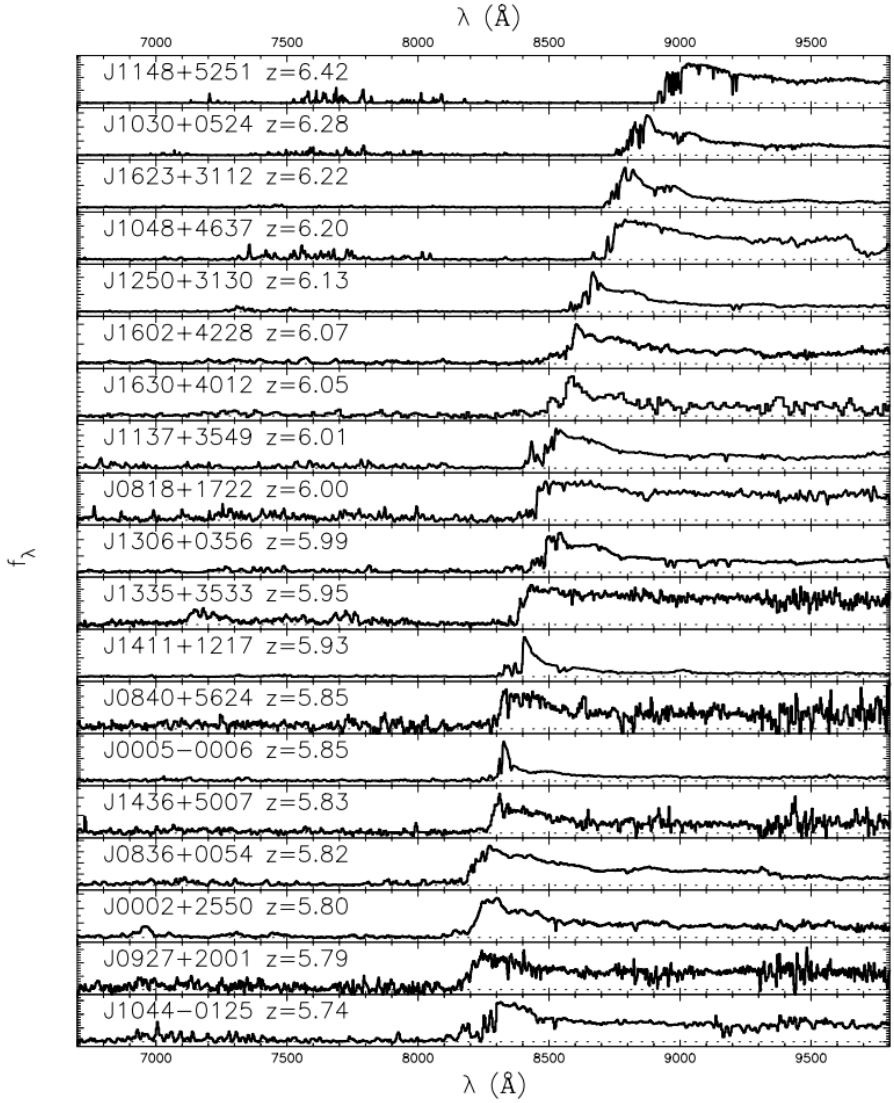

Figure II.2 – Iconic figure from Fan et al. (2006) showing 19 quasar spectra across $5.7 < z < 6.4$ displaying increasingly strong absorption. Gunn-Peterson troughs are obvious in quasars at $z > 6.1$ and occur in a few of the $z < 6.1$ objects as well.

density (log $N \lesssim 10^{16}$) absorbing clouds is distinct from that of galaxies in both occurrence rates and clustering. They serve as nearly linear tracers of density, and are likely associated with the filamentary structure of the Cosmic Web. Because of this property, the Lyman-α forest has been used as a probe of numerous cosmic properties, from the nature of dark matter (e.g., Viel et al. 2005), to measurements of the baryonic acoustic oscillations at later times (c.f. Bautista et al. 2017), to the Jeans smoothing of baryons in the IGM across cosmic time (Rorai et al., 2013).

Photo-ionisation of Lyman-α forest gas predicts a tight relation between the optical depth of Lyman-α absorption and density following reionisation. Writing the local matter over-density as $\Delta = \rho/\langle\rho\rangle$ the optical can be expressed as,

$$\tau = \tau_0 \Delta^2, \tag{II.20}$$

where τ_0 is the Gunn-Peterson optical depth given in Eq. II.19. τ_0 can be re-formulated to depend on temperature via $\tau_0 \propto \alpha(T)$, the hydrogen recombination coefficient, which varies as $\alpha \propto T^{-0.7}$ when $T \sim 10^4$K (Verner & Ferland, 1996). Meanwhile, analytical models and simulations have robustly predicted that temperature should correlate tightly with over-density following a power law $T = T_0\Delta^{\gamma-1}$, where T_0 is the temperature at mean density (e.g., Hui & Gnedin 1997; Theuns et al. 1998). Combining these effects, we arrive at the well-known optical depth–density relation,

$$\tau \propto \Delta^2 T^{-0.7} \propto \Delta^{2-0.7(\gamma-1)}. \tag{II.21}$$

Models predict that adiabatic expansion (collapse) of low (high) density regions should lead to a temperature-density relation with $\gamma > 1$ i.e. denser absorbers are hotter. Some measurements have reported an inverted relationship with $\gamma < 1$ (see e.g., Becker et al. 2007; Viel et al. 2009; Calura et al. 2012). These measurements have generally been based on the Lyman-α transmitted flux probability density function (PDF), which is sensitive to very low-density gas. The fact that the PDF prefers $\gamma < 1$ may indicate an extra source of heating at low densities, or a spatially inhomogeneous temperature field (Garzilli et al., 2012; Rorai et al., 2017).

As $z \sim 6$ is approached however, the individual Lyman-α forest components begin to blend, complicating the analysis of the forest. In addition to the study of the Lyman-α forest discussed in Section I.4.1, four techniques have thus far been successful in analysing the residual Lyman-α transmission despite its great opacity: IGM damping wing absorption, the proximity zone effect, dark pixel statistics (or dark gaps) and residual transmission spike statistics. Here I give a (very) brief overview of each in turn, as those techniques are relevant to the work presented in Chapters 3 and 5 where more technical details will be introduced.

If the Lyman-α absorption is sufficiently strong, it is possible for the damping wing of the Gunn-Peterson trough to extend significantly bluewards of the Lyman-α forest (Miralda-Escudé, 1998). This effect is called a *Gunn-Peterson damping wing*. The damping wing can extend up to ~ 30Å to the red side of the Lyman-α emission in the rest-frame. Because this feature is smooth, the expected observational effect is a weakening of the Lyman-α emission line of quasars with increasing redshift. The strength of the effect depends (at least mildly) on the topology of reionisation and strongly on the evolution of the UVB. Multiple detections of the Gunn-Peterson damping wing have been claimed (Mesinger & Haiman, 2004, 2007; Mortlock et al., 2011; Bolton et al., 2011; Schroeder et al., 2013) although the conclusions drawn for $x_{\rm H\,I}$ are necessarily model-dependent. Two important caveats of this technique are **(i)** it depends crucially on the modelling of the underlying quasar Lyman-α emission line and assumes no evolution of intrinsic quasar properties with redshift; **(ii)** it cannot distinguish between neutral gas being present in a neutral IGM or in a DLA located close in front of the quasar. Using a prior on the probability of occurrence of DLAs, Schroeder et al. (2013) claim the strongest lower limit yet at $z \sim 6$: $x_{\rm H\,I} \gtrsim 0.1$ (95% confidence). A strong GP damping wing has also been reported towards the $z = 7$ quasar J1120 (Mortlock et al., 2011; Simcoe et al., 2012; Greig et al., 2017), a claim which we will analyse in Chapter 3.

Another feature occurring close in wavelength to the Lyman-α emission line is the *proximity zone*. These regions of increased transmission can be seen just blueward of the Lyman-α emission line, right before strong absorption by the ambient IGM begins. Proximity zones are produced as UV radiation from the quasar itself ionises its immediate surroundings even within an otherwise significantly neutral IGM. Proximity zones have been detected in quasars at $z > 5.5$, and their size decreases significantly over $5.7 < z < 6.4$ albeit with very large

scatter (Carilli et al., 2010; Calverley et al., 2011; Eilers et al., 2017). Other factors besides the neutral fraction of the IGM influence the size of a proximity zone, such as the quasar lifetime, the mass of its host halo, the number of neighbouring galaxies, and the viewing angle (Bolton & Haehnelt, 2007a; Lidz et al., 2007; Wyithe et al., 2008; Maselli et al., 2009; Keating et al., 2015). In addition to the difficulties of disentangling these effects, proximity zones might not even be good tracers of the H II region around the quasar. Instead, proximity zone individual lines of sight may be terminated stochastically by self-shielded gas in small haloes located within the inhomogeneous H II bubble (Maselli et al., 2007; Keating et al., 2015).

A more model-independent measure of $x_{\mathrm{H\,I}}$ has been computed using *dark gaps* (Croft, 1998; Gallerani et al., 2008; McGreer et al., 2011, 2015). All pixels in a quasar spectrum which are 'dark' in both Lyman-α and Lyman-β are counted, their size converted to comoving length, and hydrogen column densities sufficient for total absorption are assigned to them. The resulting limit on $x_{\mathrm{H\,I}}$ is necessarily very conservative, as all absorption is assumed to be caused by the IGM instead of isolated pockets of neutral gas. As data quality improves in both SNR and resolution, the limits computed from a particular quasar will become *more* stringent (see also Greig & Mesinger 2017). The latest constraints using this method are given by McGreer et al. (2015): $x_{\mathrm{H\,I}} < 0.04+0.05$ at $z = 5.6$ and $x_{\mathrm{H\,I}} < 0.06+0.05$ at $z = 5.9$, a strong indication that reionisation has completed by $z \sim 6$. Note the conflict between this measurement and the one obtained from damping wings above.

Finally, another technique to utilise Lyman transmission is the characterisation of residual transmission spikes. At $z \sim 6$, the Lyman-α forest presents as a series of small transition spikes within the Gunn-Peterson trough (e.g., Oh & Furlanetto 2005; Becker et al. 2015b). The constraints provided by those spikes are weak, as in common models transmission peaks tends to arise from low-density voids. For instance, a single peak in a very saturated trough at $z = 7$ leads to constraints as weak as $x_{\mathrm{H\,I}} > 10^{-4}$ at $z = 6.5$ (Barnett et al., 2017). Recently attempts have been made to compare transmission spike statistics to numerical models, although this process is intrinsically very model-dependent and slightly controversial (e.g., Davies et al. 2017; Chardin et al. 2017a).

II.4 Intervening metal absorption

Atomic transitions in ions other than hydrogen and helium are referred to as *metal lines*. In a way similar to the one described above, each of these transitions also gives rise to a corresponding 'forest' of absorption lines in quasar spectra, albeit far more sparse than those created by hydrogen and helium. When these absorption lines occur at wavelengths which overlap with the Lyman-α forest, they can sometimes be a nuisance as they get confused with the hydrogen lines (e.g., Wilczynska et al. 2015). However, when correctly identified, these lines provide powerful probes of cosmic metal enrichment and galaxy formation. Here I briefly review the physics relevant to the analysis of metal lines.

The statistics which can be obtained from metal lines come from either counting the number of lines, n, over an interval, or from their distributions of either equivalent width W or column density N. Values for dN/dz and dW/dz and their evolution across cosmic time are often considered, but comparisons of measurements across very large redshift intervals are complicated by the fact that the Hubble parameter varies as

$$H^2(z) = H_0^2 \left[\Omega_{\mathrm{mat}}(1+z)^3 + \Omega_\Lambda\right], \tag{II.22}$$

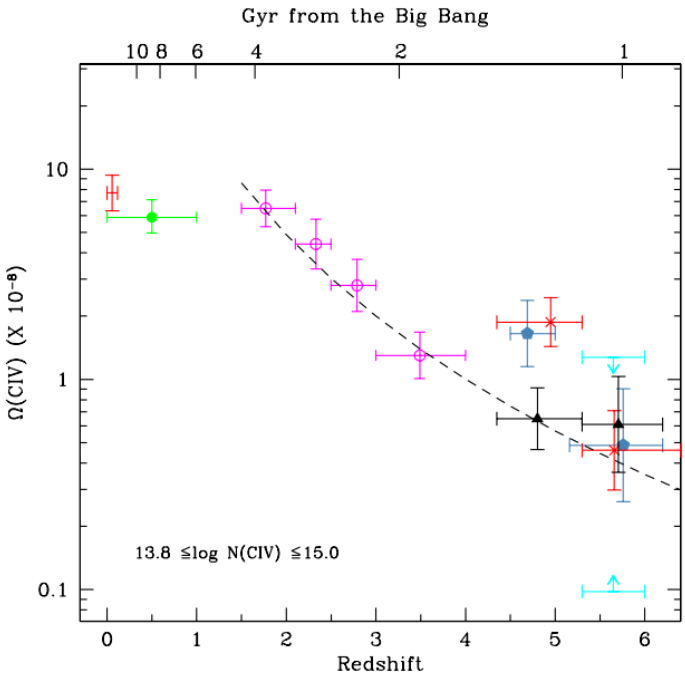

Figure II.3 – The evolution of the cosmic mass fraction of C IV contained in strong systems ($\log N \geq 13.8$), compiled from surveys across $0 < z < 6.5$. Reproduced from D'Odorico et al. (2013).

when curvature and radiation density are neglected. Number densities are often given with respect to the *absorption pathlength* X(Bahcall & Peebles, 1969):

$$\frac{dX}{dz} = (1+z)^2 \frac{H_0}{H(z)}. \tag{II.23}$$

The absorption pathlength ensures that a non-evolving population with constant comoving spatial density and physical cross-section has a constant line-of-sight density per unit X, which is not true per unit z. Like redshift, X is dimensionless.

Another common measurement is the *total mass density* of gas in a particular ionisation state. This is expressed as a fraction of the critical density ρ_{crit},

$$\Omega = \frac{\rho_{\text{ion}}}{\rho_{\text{crit}}} = \frac{H_0 m_{\text{ion}}}{c\rho_{\text{crit}}} \frac{\int N dN}{\Delta X}. \tag{II.24}$$

The critical density for a flat Universe is given by

$$\rho_{\text{crit}} = \frac{3H^2(z)}{8\pi G}. \tag{II.25}$$

The mass fraction Ω_{ion} is usually defined as a fraction of the critical density at $z = 0$, $\rho_{0,\,\text{crit}} = 1.88 \cdot 10^{-29} h^2 \text{g cm}^{-3}$ (Planck Collaboration et al., 2016b). Note that only the mass fractions corresponding to particular ions, not the elements themselves, are directly measurable. Often substantial ionisation corrections are needed to estimate the total elemental densities.

The column density (or equivalent width) distribution of intervening metal lines is parametrised by a distribution function $f(N)$,

$$\Omega = \frac{H_0 m}{c\rho_{\text{crit}}} \int_{N_{\text{min}}}^{N_{\text{max}}} N f(N) dN. \tag{II.26}$$

$f(N)$ is often approximated as a power law, an exponential function, a Schechter function or a piecewise combination of those. Depending on the slope of the chosen distribution, this often requires that a threshold N_{min} be chosen to avoid divergences. Physically, this acknowledges the fact that the smallest absorbers are produced via different physical processes than typical absorbers (such as fragmentation) which are beyond reach of current observations and simulations. The chosen value of N_{min} can have a large influence on the deduced Ω. For this reason, and especially when only a few observations are available, Equation II.26 is often approximated as,

$$\Omega = \frac{H_0 m}{c\rho_{\text{crit}} \Delta X} \sum_{\text{detections}} N_i. \tag{II.27}$$

In the limit of a large number of detections and an appropriate parametrisation, these two measures are expected to agree. In either case, however, correctly accounting for survey incompleteness can be critical.

II.4.1 Detection methods

Due to cosmological redshift, $z + 1 = \lambda_{\text{observed}}/\lambda_{\text{emitted}}$, there is a degeneracy between the wavelength of the absorbing transition and its redshift. For instance, a feature observed at $\lambda_{\text{obs}} = 10417$Å could be due to O I 1302 Å absorption at $z = 7.0$ or equally to Si II 1526 Å absorption at $z = 5.8$. To break this degeneracy, the detection and identification of intervening metal lines relies on *multiplets* and *co-occurring species*. Both of these produce intrinsic signatures consisting of more than one wavelength that have the same ratio regardless of redshift.

Multiplets arise through splitting of the energy levels between which a transition occurs. Instead of the electron simply jumping from a high energy level to a lower one, the upper level is split by the magnetic field of the nucleus via the Zeeman effect acting on electron spin. The simplest case is that of an electron jumping to or from an atomic P orbital where it is alone, which results in a two-way splitting. This produces a doublet transition with ΔE values very close to each other. Some of the most famous are Na I 5890 Å and 5896 Å, C IV 1548 Å and 1550 Å, and Mg II 2796.3 Å and 2803.5 Å. Alternatively, atoms with more complex electronic structures undergo transitions corresponding to different electronic jumps which happen to lie close in ΔE just due to the large number of possible transitions. For instance Fe II can produce dozen of transitions, some of the most useful of which are at 2344 Å, 2382 Å, 2586 Å and 2600 Å. These multiplet transitions are the easiest intervening metal lines to identify for a number of reasons. First, since all transitions originate from the same gas, the Voigt profiles of the lines must share the same parameters, in particular the Doppler width related to temperature. For the same reason, the intrinsic velocity offset between multiple components must be exactly zero. If any of those parameters are not consistent within the measurement errors, the absorption feature can be safely discarded as a chance alignment of unrelated transitions. Secondly, the relative strengths of the multiplet components are set by the transition cross-sections. For instance in p-orbital doublets the transition with lower ΔE is twice as likely as its pair. Again

the measured strengths of absorption must be consistent with this, as any deviation signals that either the features are unrelated, contaminated by interlopers, or that the absorber is only partially covering the line of sight (which can generally only happen for absorbers close to the background source).

Multiple related transitions are also seen from ions which often occur together and whose ionisation potentials are similar. For instance, carbon and silicon are both common products of stellar enrichment processes and their transitions C II 1334 Å and Si II 1526.71 Å have similar ionisation potentials ϵ.[3] They can therefore often be found together and can be used in a comparable manner to multiplets in the detection of intervening objects, with a few important differences. The relative strengths of these transitions depend on the column densities of both elements and are therefore not fixed; one could potentially be present without the other. Secondly, the absorption can occur in different sub-regions of the absorber with different temperatures and velocity offsets. Finally, which elements co-occur depends strongly on the local ionising potential. In practice, sub-regions of the absorber which display transitions with large ϵ (e.g., C IV, Si IV, N V) and those with low ϵ (e.g., O I, C II, Si II, Mg I) can be spatially separated by being exposed to different radiation. If the absorption is co-spatial, then the relative abundance of different ions of the same element (e.g., Si II and Si IV) can be used to measure the ionization of the gas.

Once a system has been detected at a redshift $z_{\rm abs}$ via the presence of multiple lines, single transition lines located at exactly $z_{\rm abs}$ can be searched for and associated with the absorber. Such alignments are extremely unlikely to occur by chance due to the small tolerance in Δv between co-occurring species, but it is important to quantify exactly how many systems are being missed or falsely detected. This is usually achieved through the automation of the detection procedure to remove human bias while checking that the results are in agreement with a well-versed researcher. Artificial systems are inserted into the data and their recovery is attempted to gauge the minimal sensitivity levels. Similarly, fictional doublet systems with the wrong spacings or relative strength can be inserted to ensure that they are not picked up as detections. These are only a few of the techniques used to evaluate detection completeness; more details are given in Chapter 4 where the spectrum of the $z = 7$ quasar ULAS J1120+0641 is searched for intervening absorbers.

II.5 This Thesis

The remainder of this thesis presents three approaches to exploring the first billion years of the Universe using the techniques outlined in this chapter. Chapters 3 and 4 make use of the only currently known quasar at $z > 7$, ULAS J1120+0641, observed with the GNIRS instrument and in a deep, 30h X-Shooter exposure, respectively. This quasar was discovered in 2011 by the UKIRT Infrared Deep Sky Survey (UKIDSS, Lawrence et al. 2007) using the UK Infrared Telescope. In Chapter 5, I make use of ~ 90 $z > 5.7$ quasars to analyse the statistics of the high-redshift IGM Lyman-α opacity. The scientific motivations are as follows:

1. We re-evaluate the significance of weak Lyman-α emission lines in high-z quasars. We

[3] More precisely, these transitions will generally occur in the same region because the ionising potentials I→ II and II → III are similar. Both I→ II transition energies are less than 13.6 eV, corresponding to hydrogen, and both II → III transition energies are above it. Both these species will therefore be present if the Lyman limit is even slightly optically thick.

find that quasars with weak Lyman-α do exist at low redshifts without signs of an IGM damping wing. Although they are rare overall, weak Lyman-α objects are in fact *common* among low-z quasars selected for extreme C IV blueshifts, a feature displayed by the $z = 7.08$ quasar ULAS J1120+0641.

2. We exploit the only quasar at $z > 7$ to search for intervening metal systems along its line of sight. We measure occurrence rates and cosmic mass fractions up to $z = 7$, the first time this has been done for many of these metals. Although a single line of sight is insufficient to obtain tight constraints, we find a significant and unexpected over-density of weak Mg II absorbers at $z > 6$. We also find indications of partial covering of the quasar region by an absorber located -920 km s^{-1} from the object. This could cast doubt on the interpretation of the quasar's short proximity zone, a crucial part of the argument for an IGM damping wing.

3. We assemble the largest catalog ever studied of $z > 5.7$ quasars, consisting of 89 spectra taken on 8 different spectrographs, bringing together data from the SDSS, DES and SHELLQs surveys as well as new reductions of public archival data. This constitutes the first truly statistical sample of high-z quasars, which is necessarily required in order to analyse the evolution of Lyman-α transmission.We compare our results to the outputs of a range of cosmological simulations and confirm that the distribution of Lyman-α opacities is strongly inconsistent with a uniform UVB model at $z > 5.0$.

III

Re-examining the case for neutral gas near the redshift 7 quasar ULAS J1120+0641

What you see is what you see. And what you don't see, well, you don't see it.

Nicolas Dupré

Abstract

SIGNS of damping wing absorption attenuating the Lyman-α emission line of the first known $z \sim 7$ quasar, ULAS J1120+0641(J1120), recently provided exciting evidence of a significantly neutral IGM. This long-awaited signature of reionisation was inferred, in part, from a deficit of flux in the quasar's Lyman-α emission line based on predictions from a composite of lower-redshift quasars. The composite sample was chosen based on its C IV emission line properties; however, as the original study by Mortlock et al. (2011) noted, the composite contained a slight velocity offset in C IV compared to J1120. Here we test whether this offset may be related to the predicted strength of the Lyman-α emission line. We confirm the significant (~ 10 per cent at r.m.s.) scatter in Lyman-α flux for quasars of a given C IV velocity and equivalent width found by Mortlock et al. (2011). We further find that among lower-redshift objects chosen to more closely match the C IV properties of J1120, its Lyman-α emission falls within the observed distribution of fluxes. Among lower-redshift quasars chosen to more closely match in C IV velocity and equivalent width, we find that J1120 falls within the observed distribution of Lyman-α emission line strengths. This suggests that damping wing absorption may not be present, potentially weakening the case for neutral gas around this object. Larger samples of $z >7$ quasars may therefore be needed to establish a clearer picture of the IGM neutral fraction at these redshifts.

III.1 Introduction

The reionisation of hydrogen in the intergalactic medium (IGM) is believed to coincide with the buildup of the first galaxies. Determining when and how the IGM became ionised can therefore deliver unique insights into the earliest epochs of galaxy formation. Evidence from the Lyman-α forest in high-redshift quasars indicates that the process largely finished by $z \sim 6$ (Fan et al., 2006; Becker et al., 2007). While the exact timing remains unknown the recent Planck results (Planck Collaboration et al., 2016a) favour models of late reionisation, which motivates the observational search for signs of reionisation.

The recent discovery of the most distant quasar known, ULAS J1120+0641 (hereafter J1120; Mortlock et al. 2011), with a redshift of 7.0842 $\pm$ 0.0004 (Venemans et al., 2012), provides a unique opportunity to probe the epoch of reionisation at a time when the process was potentially just ending. As noted by Mortlock et al. (2011), this object exhibits strong absorption on the blue side of its Lyman-α emission line. The red side of its Lyman-α line may also be somewhat weaker than expected based on the strength of its other emission lines, notably C IV . These traits have jointly been interpreted as evidence for neutral gas in the vicinity of J1120, with the weak Lyman-α line being due to damping wing absorption extending to $\lambda_{\rm rest} \geq 1216$ Å. In a scenario where the neutral gas is intergalactic, models of the quasar's proximity zone suggest that the surrounding IGM may be $\geq$10 per cent neutral (Bolton et al., 2011). The lack of intervening metal absorption lines further suggests that any neutral gas would have to be highly metal poor, particularly if it is confined to a discrete absorber (Simcoe et al., 2012).

The case for neutral gas in the vicinity of J1120, whether localized or intergalactic, depends strongly on the reality of the damping wing absorption. A truncated proximity zone, while unusual, could be due to an optically thick ($N_{\rm H\,I} \gtrsim 10^{17}$ cm^{-2}) absorber in the vicinity of the quasar, but a damping wing requires a large column density ($N_{\rm H\,I} \gtrsim 10^{20}$ cm^{-2}) of neutral gas. A challenge in analysing such high-redshift quasars, however, is that their intrinsic spectra, including their Lyman-α emission lines, exhibit significant variation between objects. This combined with the fact that the blue side of Lyman-α is often strongly affected by Lyman-α forest absorption (generally arising from ionised gas) means that the intrinsic shape of Lyman-α is largely unknown a priori. This makes it difficult to estimate the extent to which the emission line flux has been absorbed, and in the case of J1120, whether damping-wing absorption may be present.

One approach to resolving this issue has been to estimate the intrinsic quasar spectrum near Lyman-α using either Principal Component Analysis (PCA, Suzuki et al., 2005; Pâris et al., 2011; Greig et al., 2016) or composites made from lower-redshift objects (e.g., Vanden Berk et al., 2001; Cool et al., 2006). The general approach is to find lower-redshift quasars whose spectra match the object of interest redward of Lyman-α. The spectra of the lower-redshift objects, which are less affected by Lyman-α forest absorption, can then be used to predict the unabsorbed continuum around the Lyman-α emission line and over the forest. This was the approach taken by recent studies of J1120 (Mortlock et al., 2011; Simcoe et al., 2012). In these cases the composite was constructed primarily to match the C IV emission line, whose properties are known to correlate with those of Lyman-α (Richards et al., 2011).

The fidelity of a composite match will naturally depend on the availability of similar objects at lower redshifts. However, the C IV line of quasars exhibits a source systematic velocity shift (Richards et al., 2002; Shang et al., 2007) which is usually attributed to the presence of strong winds and jets associated with the quasar (Proga et al., 2000). In addition the Equivalent Width

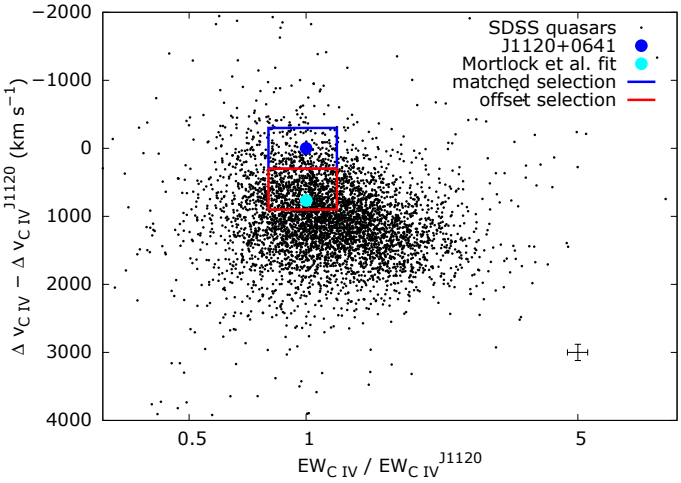

Figure III.1 – C IV Equivalent Width-Blueshift Anomaly diagram showing the C IV emission properties of 5207 SDSS quasars in black and J1120 in blue. This includes all DR7 quasars in the range $2.4 < z < 4$ for which spectra with SNR>5 were available. Error bar in the bottom right corner is representative for points located in the middle of the distribution; scatter in extreme values is due to objects whose C IV line is not well modelled by the spline fit, such as BAL quasars. Thick lines: initial cuts from which the matching (blue) and offset (red) selections were extracted (see text). The location of the composite quasar spectrum from Mortlock et al. (2011) is shown in light blue.

(W) of the C IV line is known to correlate with Equivalent Width of Lyman-α. Mortlock et al. (2011) match J1120 by selecting objects based on their C IV W and blueshift, but acknowledge that this is complicated by the fact that J1120 has a large C IV blueshift and is matched by relatively few lower-z objects. The sparsity of suitable objects is aggravated by the need to restrict the selection to objects within the narrow redshift range $2.3 \lesssim z \lesssim 2.6$ chosen to minimize the impact of Lyman-α forest absorption. A compromise thus had to be made in order to obtain a large enough sample of objects, resulting in a small misalignment in C IV blueshift between the composite and J1120's spectrum. In this paper we examine whether the difference in blueshift, although small, might be having an unanticipated effect on the predicted strength of the Lyman-α emission line.

III.2 Methods

We aim to predict the strength of the Lyman-α emission line without making use of a composite spectrum. Our approach is to select two samples of lower redshift quasars solely on their C IV emission properties, regardless of how rare they may be. One sample is designed to contain objects similar to J1120 in C IV properties, except with a small blueshift mismatch; the objects in the second sample are chosen to display no such offset. We then measure the Lyman-α flux strength for all objects, which allows us to establish the rarity of quasars with Lyman-α emission intrinsically as weak as in J1120 in both samples.

Our sample of comparison spectra is drawn from the Sloan Digital Sky Survey (SDSS) Data Release 7 (DR7) quasar catalog (York et al., 2000). The DR7 catalog contains 10871 objects

Selection	W_{CIV}/W_{CIV}^{J1120}	$\Delta v_{CIV} - \Delta v_{CIV}^{J1120}$
Matching	0.8 → 1.2 Å	-300 → 300 km s^{-1}
Offset	0.8 → 1.2 Å	300 → 900 km s^{-1}

Table III.1 – Selection criteria for our initial C IV emission line cuts. The matching selection is chosen to match the C IV emission of J1120 as well as possible. The offset selection is chosen to mismatch J1120's C IV emission line blueshift by 600 km s^{-1} or about 3 Å.

in the range $2.4 \leq z \leq 4$, where both the Lyman-α and C IV emission lines are in the observed wavelength range. DR7 was chosen because of the availability of high quality redshifts measured by Hewett & Wild (2010). In this work we use only objects where the systemic redshift was determined primarily from the C III] emission line complex. These redshifts were used to shift the spectra into the rest frame, and for assessing the relative blueshift of the C IV line.

Similarly to Mortlock et al. (2011), we selected sub-samples of quasars based on their C IV rest-frame equivalent width, $W_{\rm C\,IV}$, and velocity shift relative to systemic, $\Delta v_{\rm C\,IV}$. We do not attempt to provide an absolute value of J1120's C IV line blueshift, as doing so requires more careful treatment of the underlying continuum as well as nearby emission lines, in particular Fe II. For an up-to-date value, see De Rosa et al. (2014). Since the ultimate goal is to identify objects with properties similar to J1120, we used the J1120 C IV emission line itself as a basis for the classification. Using the FORS2+GNIRS spectrum presented by Mortlock et al. (2011), we fit the J1120 C IV line over the wavelength range $1435 < \lambda_{\rm rest} < 1640$ Å using two components, a power law fit over $1435 < \lambda_{\rm rest} < 1480$ Å and $1580 < \lambda_{\rm rest} < 1640$ Å, plus a spline, fit by hand, to the emission line flux in excess of the power law shown in Fig. III.2. We then fit each object in the DR7 catalog over the same wavelength range using a power law plus an emission line template based on the J1120 spline. The emission line fit, normalized by the power law continuum, is given by

$$F_{\rm C\,IV}(\lambda') = aF_{\rm C\,IV}^{1120}(\lambda)\,, \tag{III.1}$$

where $F_{\rm C\,IV}^{1120}$ is the J1120 emission line fit, normalized by its power law, and a is a scaling factor. The rest-frame wavelengths are related as

$$\lambda' = \lambda\left(1 + \frac{\delta v}{c}\right)\left[1 + s\frac{(\lambda - 1540\,\text{Å})}{\lambda}\right]\,, \tag{III.2}$$

where δv is the velocity shift of the C IV line relative to J1120, and s is a stretch factor about 1540 Å, which lines near the peak of the C IV line in J1120. Nominally, objects with C IV lines that are well matched to J1120 will have $a \simeq 1$, $\delta v \simeq 0$, and $s \simeq 0$. Given these parameters, $W_{\rm C\,IV}$ for each object as a fraction of that of J1120 can then be recovered by $W_{\rm C\,IV}/W_{\rm C\,IV}^{1120} = a(1+s)$.

Fig. III.1 shows the resulting C IV emission parameter space for 7025 SDSS quasars over $2.4 \leq z \leq 4.0$ for which spectra with Hewett & Wild (2010) C III] redshifts were available. As noted by Mortlock et al. (2011), the J1120 C IV emission line appears to be highly blueshifted compared to the general population of quasars. This makes it fall at the edge of the usual distribution of C IV parameters.

We proceeded to select one sample of objects located close to J1120 in C IV parameter space, and another sample with the C IV blueshift offset from the nominal value by ∼3 Å (∼600 km s^{-1}). This offset is chosen to mimic the apparent offset in C IV between the J1120

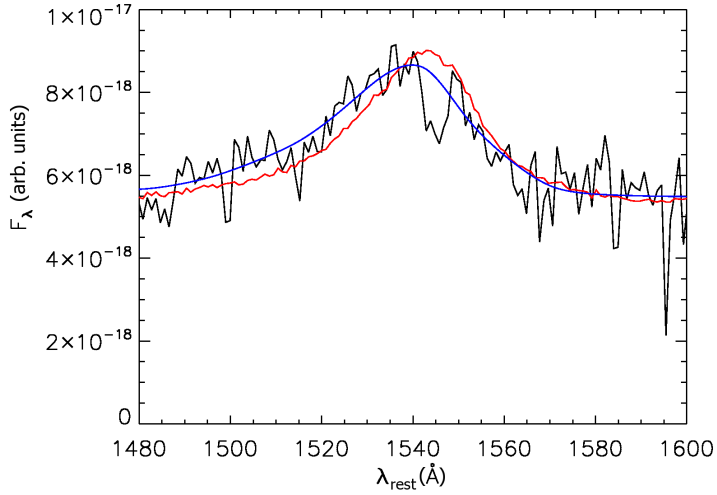

Figure III.2 – Black: C IV emission line of J1120 from GNIRS spectrum. Red: Composite quasar spectrum from Mortlock et al. (2011). Blue: Spline fit to J1120's C IV emission line. The distance between the peaks of the fits is ∼3 Å in the rest frame of the quasar, corresponding to a velocity offset of ∼600 km s^{-1}.

spectrum and the composite fit of Mortlock et al. (2011), as shown in Fig. III.2. The cuts we used are shown in Table III.1.

A cut of SNR ≥ 5 per pixel, measured by making use of the associated SDSS error arrays over the wavelength range 1450 - 1500 Å, was also applied in the interest of measurement accuracy. For comparison the GNIRS spectrum of J1120 has SNR$_{\rm J1120}$ ∼14 per velocity bin over the same range. This initial cut yielded unrefined samples consisting of 204 C IV-matched quasars and 709 offset quasars. Due to the large uncertainties in the values of W and blueshift for the objects in our sample, we expected the initial cuts to be inefficient at finding all of, and only, excellent matches for J1120's C IV emission. For this reason the initial selection boxes are large and a second, manual cut was necessary to refine the selection.

Next, each object's C IV emission line was overlayed with J1120's C IV emission line and similarity was evaluated visually. Objects which displayed too much or too little line asymmetry, as well as those whose C IV emission peak was visually offset from J1120's corresponding value, were rejected and excluded from the sample. This yielded a refined sample of 111 objects which closely matched J1120's C IV emission. The same process was repeated for the C IV-offset sample, yielding a refined sample of 216 objects with a C IV blueshift offset. The mean redshifts for the matched and offset selections were $z = 2.877$ and $z = 2.912$ respectively while the median values were $z = 2.898$ and $z = 2.937$ respectively. A two-sided KS test revealed no evidence that the two samples had been drawn from different distributions (p>0.10).

Measuring the intrinsic Lyman-α flux is particularly difficult for our samples since they include quasars with redshifts as high as $z = 4$, for which Lyman-α forest absorption is a greater issue. This is further complicated by the complex shape of the Lyman-α emission line over this range, which includes contributions from asymmetric Lyman-α emission as well as N V. For these reasons the mean unabsorbed continuum in each object was estimated over the range 1216-1220 Å, which was chosen to maximize our sensitivity to the Lyman-α emission

strength, while minimizing the impact of Lyman-α forest absorption. Even in this wavelength range, however, Lyman-α and metal absorption lines can potentially obscure the intrinsic continuum, particularly at higher redshifts. We therefore inspect each object visually and estimate the continuum by eye, accounting for absorption lines to the extend possible. This will naturally be an imperfect process; however, the errors in the continuum estimate arising from absorption lines are expected to impact both samples equally. To mitigate any bias caused by manual measurement, we measured the two samples together with the objects in random order. As a further check, we calculate the mean spectrum of the two samples without attempting to correct for absorption lines (see below).

Emission line fluxes for each object were calculated as a fraction of the underlying power-law continuum. The flux of each object was normalised by measuring the mean flux over the wavelength range 1450 - 1500 Å and the continuum emission was modeled by a power-law fitted iteratively over narrow, emission line free regions of the spectrum. These fitting windows spanned the rest wavelength ranges 1320 to 1350 Å, 1435 to 1480 Å and 1600 to 1700 Å. The range 2000 - 2050 Å was used in addition for $z \leq 3$ objects for which those wavelengths were present in the rest frame. The peak flux over C IV was measured by the same method for all objects as well as J1120 to check that the C IV emission was well matched by our selections and to ensure that direct comparison with the flux value over the Lyman-α red wing was meaningful.

III.3 Results and Discussion

Among the objects which matched the C IV emission in shape and W but which were offset in blueshift by $\sim$ 600 km s^{-1}, only three per cent of objects showed Lyman-α emission weaker than J1120 over the same range (Fig. III.3). This is consistent with the assessment by Mortlock et al. (2011) that J1120 is a $\sim 2\sigma$ outlier from the composite they used. In contrast, among the selection which matched J1120 in C IV emission, we find that 31 per cent have Lyman-α emission over the red wing weaker in magnitude than J1120 (Fig. III.4). In addition our C IV measurements are shown to demonstrate no residual correlation in these small samples ($R_{\rm matched} = 0.20$, $R_{\rm offset} = 0.07$), demonstrating that the C IV emission line of J1120 is typical among these samples. We therefore find no compelling evidence that the Lyman-α flux of J1120 is anomalously low.

The averages and 68 (95) per cent bounds in Lyman-α flux values for the matched and offset sample are $F_{L\alpha} = 2.64^{+0.29}_{-0.32}\,(^{+0.63}_{-0.50})$ and $F_{L\alpha} = 2.88^{+0.17}_{-0.32}\,(^{+0.60}_{-0.52})$ respectively, corresponding to variations of $\sim \pm 10$ ($\sim \pm$ 20) per cent in both distributions. These values are similar to the corresponding value of 13 per cent quoted by Mortlock et al. (2011) for the spread in Lyman-α flux once C IV blueshift is constrained; however the distributions are mildly non-gaussian since they have a tail and those numbers are given as a consistency check only. In addition these values of the spread are given at a particular wavelength.

Small sample size, low SNR and scatter mean that our sample might not be suitable for making an appropriate composite fit to J1120, and such an analysis is beyond the scope of this work. Fig. III.5 presents the average spectrum for both our selections; those were made by interpolating the spectra onto a common wavelength array and taking the mean value pixel-by-pixel. The spectra were power-law corrected and normalised but absorption was not addressed. The difference in flux between the peaks of the average spectra in Fig. III.5 is comparable to

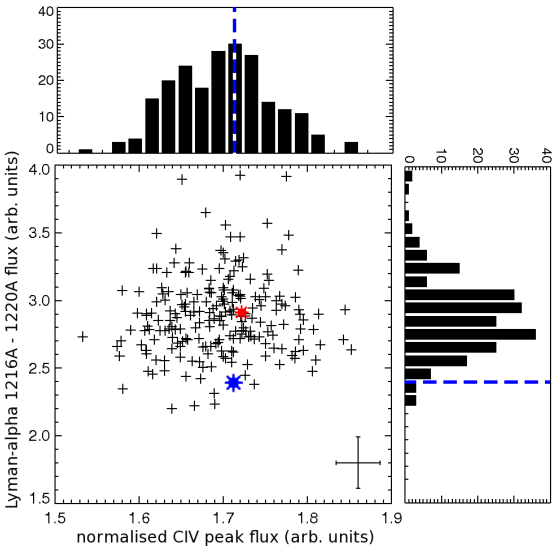

Figure III.3 – Comparison of C IV peak flux and Lyman-α red wing flux for 216 quasars matching J1120 in C IV W and offset from J1120's correct C IV Blueshift value by $\sim$ 600 km s^{-1}, drawn from the red box in Fig. III.1. J1120's location is shown by a thick blue star. Red star indicates the location of the Mortlock et al. (2011) composite quasar spectrum when measured in the same way. The Lyman-α red wing flux is measured between 1216Å– 1220Å. Among this population J1120 is a 97 per cent outlier in Lyman-α red wing flux. Representative error bar shown in the bottom right corner.

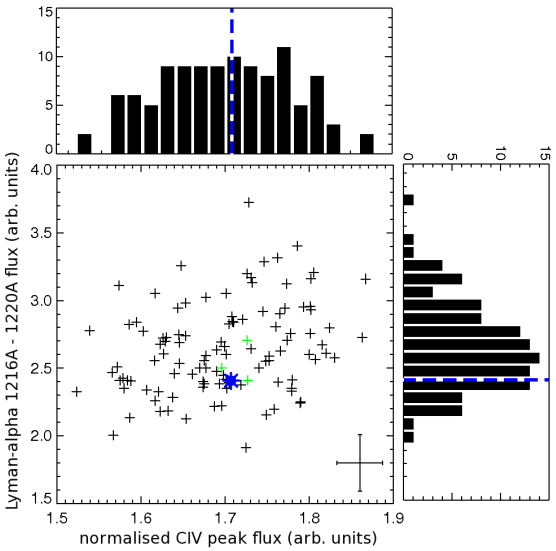

Figure III.4 – Comparison of C IV peak flux and Lyman-α red wing flux for 111 quasars matching J1120's C IV emission in both W and Blueshift, drawn from the blue box in Fig. III.1. J1120's location is shown by a thick blue star. The objects highlighted in green are shown in Fig. III.6. The Lyman-α red wing flux is measured between 1216Å– 1220Å. Among this population, 31 per cent of objects have Lyman-α lines that are weaker than J1120. Representative error bar shown in the bottom right corner.

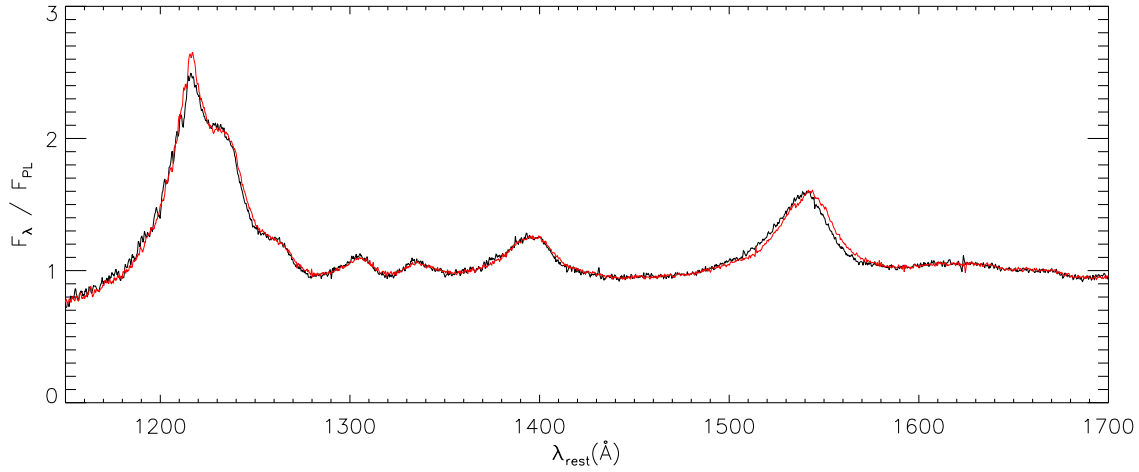

Figure III.5 – Average spectra of the matched selection of quasars (black) and the offset selection (red). Note how a small mismatch in the C IV emission line leads to large variation of the Lyman-α peak flux, while the low ionisation lines are not affected. The spectra were normalized by dividing by a fitted power-law; absorption was not taken into account and as a consequence these averages do not represent the underlying continuum emission. Uncorrected absorption tends to both lower and smooth the peaks hence the difference in mean spectra is less pronounced than in the ensemble distributions (see Fig. III.3 and III.4).

the difference between the averages of the distributions in Fig. III.3 and III.4; however the effect of absorption is to weaken the emission feature in the average spectra. This confirms the fact that a small mismatch in the C IV emission line can have a significant effect on Lyman-α emission. Although this effect is not large, it is sufficient to significantly impact the degree to which J1120 is an outlier. The amount by which the average spectra differ in C IV blueshift is similar to the mismatch between J1120 and its composite fit as published in Mortlock et al. (2011), condoning our use of the value of a ~ 600 km s^{-1} C IV blueshift mismatch.

Example objects matching J1120's spectrum extremely well over its entire continuum and down to the onset of the Lyman-α forest at 1216Å are shown in Fig. III.6 along with their respective redshifts. Using the strength of flux right of 1216Å in a manner similar to J1120, one would thus falsely infer the presence of a damping wing in those objects. A large variety of Lyman-α peak shapes exists among the objects which match J1120's C IV line; in particular the N V emission line at 1240Å overlaps with Lyman-α in a non-trivial way especially when the latter is weak. We make no attempt to constrain the N V emission. The majority of objects in our sample display a visible Si II emission line at 1260Å, suggesting that the high-ionisation lines C IV, N V and Lyman-α are blueshifted together while low-ionisation lines Si II, O I, C II are not. Our findings are in agreement with Richards et al. (2011) who identify a link between a large NV/Lyman-α ratio in quasars with extreme C IV blueshifts as well as a possible C IV blueshift/Lyman-α blueshift correlation. This effect also helps to resolve the apparent mismatch of N V in the Mortlock et al. (2011) composite.

III.4 Caveats

We have argued that the scatter in Lyman-α flux for a given C IV W, combined with a small decrease in Lyman-α flux with increasing C IV blueshift, is sufficient to bring the red wing of the Lyman-α line of J1120 within the distribution seen in lower-redshift objects. Reconstructing the unabsorbed continuum of a quasar is known to be a challenging problem, however (e.g., Pâris et al., 2011, and references therein), and it is important to bear in mind certain caveats. For example, accurate measurements of C IV blueshifts require reliable estimates of the systemic quasar redshifts. For the comparison SDSS spectra we have used redshifts based on C III] from Hewett & Wild (2010), assuming that these are independent from the C IV and Lyman-α properties. The C III] feature is a blend of Al III, Si III, and C III] emission, however (e.g., Vanden Berk et al., 2001), and so the measured redshift will depend on the relative strengths and velocity shifts of these components. If the properties of this emission complex correlate with the properties of the C IV and/or Lyman-α lines, then this may introduce systematic biases into the relationship we infer between C IV and Lyman-α.

A robust reconstruction of the intrinsic Lyman-α flux may ultimately require a more sophisticated analysis than the one presented here. We have shown that there is some uncertainty in the Lyman-α profile related to the C IV properties, and hence motivation to re-examine the claim of a damping wing. A conclusive analysis, however, may need to take into account the detailed properties of all available emission lines, and be verified through tests on lower-redshift quasars with known Lyman-α profiles.

Recent work by Greig et al. (2017) has attempted such a reconstruction of the underlying continuum of J1120 through Principal Component Analysis (PCA) of low redshift quasars ($2.08 < z < 2.5$). The deconstruction into a minimal number of orthogonal principal components aims to reveal correlations between emission properties independently of a physical model. While the method of Greig et al. (2017) does succeed at reconstructing continua bluewards of Lyman-α in their sample of low-redshift quasars, we note that the region of parameter space in C IV emission properties spanned by their training sample excludes objects such as the ones presented here. It is therefore unclear whether the correlations involved in the emergence of such extreme C IV emission lines are accurately captured by their final set of components. Moreover, we draw attention to recent hints that quasars at $z > 6.5$ might show extreme C IV line blueshifts as the rule rather than the exception (Mazzucchelli et al., 2017). If new physical processes start manifesting in this early regime, those would not be reflected in components derived from a sample of quasars at lower redshift.

As noted earlier, the J1120 redshift is measured from [C II] 158 μm emission from the ISM of the host galaxy (Venemans et al., 2012). There is some indication, however, that C III] in J1120 is slightly blueshifted with respect to [C II] (see Fig. 1 of Mortlock et al., 2011). If this is generally true for quasars with large C IV blueshifts, then we should potentially adopt the C III] redshift for J1120 when comparing it to objects with redshifts derived from C III], as noted by Greig et al. (2017). This would give a smaller relative blueshift for C IV in J1120, meaning that our comparison objects should be drawn from a region closer to the red box in Fig. III.1. By the arguments presented here, this would imply a somewhat higher expected intrinsic Lyman-α flux over 1216-1220 Å. On the other hand, the observed wavelength at which discrete Lyman-α forest absorption begins in J1120, i.e., the shortest wavelength at which a smooth damping could be detected, would correspond to a longer rest-frame wavelength, and hence a weaker portion of the Lyman-α emission line profile. It is therefore unclear whether adopting a lower

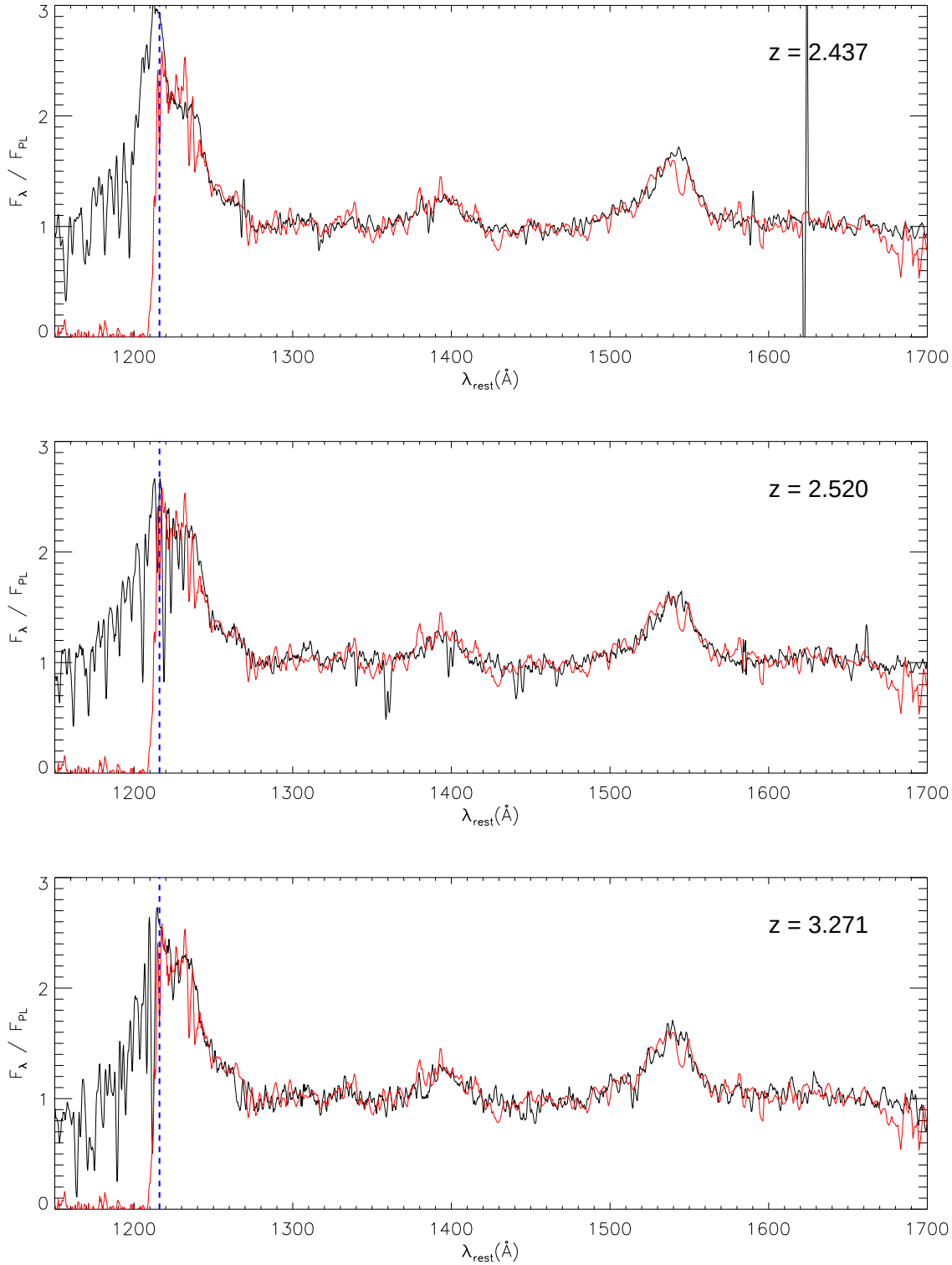

Figure III.6 – SDSS spectra of objects at lower redshifts which match J1120+0641's spectrum in C IV as well all the way to the onset of the Lyman-α forest at 1216Å, indicated by the blue dashed line. The GNIRS spectrum of J1120, binned in bins of 3, is overlayed in red. All spectra are normalized by a fitted power-law. SDSS spectra are binned in bins of 5. The C IV parameters of these objects are shown as green dots on Fig. III.4.

redshift for J1120 would strengthen the case for damping wing absorption.

III.5 Conclusions

Composites quasar spectra have been used in the past to study the IGM around J1120 and hypothesize a large fraction of neutral gas or very metal-poor gas. Those claims were based on J1120's seemingly highly-absorbed Lyman-α emission (Mortlock et al. (2011) using the quasar composite technique from Hewett & Wild 2010; Simcoe et al. 2012; Bolton et al. 2011). However those authors acknowledge the difficulty of producing a composite fit to J1120, due primarily to the rarity of suitable objects. As a result, composite fits have mismatched its C IV emission line by ~ 600 km s^{-1} or more in blueshift. Mortlock et al. (2011) suggested that this mismatch might have an impact on the existence of a Lyman-α damping wing.

We have tested the effect of this small C IV mismatch by selecting a sample of 111 SDSS DR7 quasars at $2.4 < z < 4$ that more closely match J1120's C IV emission, and a second sample of 216 quasars whose C IV emission lines match J1120's in W but are offset blueward by $\sim$600 km s^{-1} ($\sim$3Å in the rest frame). We find that among a population of quasars which match its C IV emission in equivalent width and shape, but offset by ~ 600 km s^{-1} from the correct C IV line systematic blueshift, J1120 appears anomalous to a confidence of 97 per cent, in agreement with Mortlock et al. (2011) who find J1120 to be a 2σ outlier from their composite fit. When compared to a population of lower-redshift quasars which match its C IV line emission, however, J1120's Lyman-α flux does not appear to be anomalously weak.

In light of our results we suggest that the shape of the Lyman-α line in J1120 is not obviously due to a damping wing. We note that the case for neutral gas in the vicinity of J1120 depended on both the damping wing absorption and the shortness of the proximity zone (Bolton et al., 2011). The latter could be due to an optically thick absorber near the quasar redshift, a point we will address in Section IV.4.6. Independent of the ULAS J1120+0641 constraints, however, there is potential evidence for a significantly neutral IGM at redshifts $z \geq 7$ from the evolution of the Lyman-α emission fraction in galaxies (e.g., Caruana et al., 2012; Treu et al., 2013; Tilvi et al., 2014; Choudhury et al., 2009). Direct evidence from quasars would significantly strengthen this claim. Larger samples of quasars at $z \geq$7, along with a more robust method to reconstruct their intrinsic Lyman-α flux, are therefore of great interest for future reionisation studies.

IV

A deep search for metals near redshift 7: the line-of-sight towards ULAS J1120+0641

A long time ago, in a galaxy far, far away...

George Lucas

Abstract

We present a search for metal absorption line systems at the highest redshifts to date using a deep (30h) VLT/X-Shooter spectrum of the $z = 7.084$ QSO ULAS J1120+0641. We detect seven intervening systems at $z > 5.5$, with the highest-redshift system being a C IV absorber at $z = 6.51$. We find tentative evidence that the mass density of C IV continues to decline with redshift, while the number density of C II systems remains relatively flat over $6 < z < 7$. These trends are broadly consistent with models of chemical enrichment by star formation-driven winds that include a softening of the ultraviolet background towards higher redshifts. We find a larger number of weak ($W_{\rm rest} < 0.3$ Å) Mg II systems over $5.9 < z < 7.0$ than predicted by a power-law fit to the number density of stronger systems. This is consistent with trends in the number density of weak Mg II systems at $z \lesssim 2.5$, and suggests that the mechanisms that create these absorbers are already in place at $z \sim 7$. Finally, we investigate the associated narrow Si IV, C IV, and N V absorbers located near the QSO redshift, and find that at least one component shows evidence of partial covering of the continuum source.

IV.1 Introduction

High-redshift quasi-stellar objects (QSOs) are powerful and versatile probes for studying the both intergalactic medium (IGM) and the circum-galactic media around galaxies near the epoch of reionisation. The study of Lyman-α transmission along the line-of-sight to QSOs has revealed a rapidly evolving IGM at $z \sim 6$ (e.g., Fan et al. 2006, Becker et al. 2015b), suggesting that hydrogen reionisation may be in its final stages near that redshift (e.g., Gnedin & Fan 2006, Mesinger 2010, Gnedin et al. 2016). Meanwhile, QSO near zones offer a valuable probe of the ultraviolet background (UVB) up to at least $z \sim 6$ (Wyithe et al. 2014, Bolton & Haehnelt 2007a, Maselli et al. 2009, Carilli et al. 2010). The proximity zone of the highest-redshift known QSO, ULAS J1120+0641 ($z = 7.0842$), has even provided hints of a partially neutral IGM and/or chemically pristine circum-galactic gas at $z \sim 7$ (Mortlock et al. 2011, Simcoe et al. 2012, Bolton et al. 2011, Greig et al. 2016; but see Chapter III).

Metal absorption lines tend to trace the metal-enriched gas located around galaxies, and thus probe a different mass and density regime than the Lyman-α forest and near zones. Metal lines provide insight into star formation and feedback processes, and also offer a means to study galaxies that are too faint to detect in emission. In addition, the elemental abundances in metal absorbers provide crucial information on the nature of the earliest stellar populations.

A number of studies have have now traced metal enrichment out to $z \sim 6$ using both highly ionized (C IV, Si IV) and low-ionization (e.g., C II, O I, and Mg II) species (Ryan-Weber et al. 2006, 2009; Simcoe 2006; Simcoe et al. 2011; Becker et al. 2009, 2011; Matejek & Simcoe 2012; Matejek et al. 2013; D'Odorico et al. 2013). For Mg II, the redshift frontier has recently been pushed back to $z = 7$ (Chen et al., 2016). Meanwhile, numerical simulations have explored the effects of varying star formation histories, early galaxy feedback mechanisms, and the effect of a declining ultraviolet background (UVB) on the number density and ionisation state of metal absorbers (e.g., Oppenheimer et al. 2009, Finlator et al. 2015, Keating et al. 2016). A recent review of this subject can be found in Becker et al. (2015a).

The incidence rate of both highly ionized and low-ionization species is observed to evolve with redshift, although the rate of evolution is still somewhat unclear. For example, the incidence rate of C IV appears to decrease from $z = 1.5$ to $z = 3.5$ and may decrease faster with redshift above $z = 4.5$, falling by a factor of ~ 10 over the entire range (D'Odorico et al., 2013). This has been interpreted as the result of ongoing carbon enrichment in the vicinity of galaxies due to outflows (e.g., Oppenheimer & Davé 2006) as well as a possible softening of the UVB towards higher redshift, which impacts the ionization state of carbon (Oppenheimer et al. 2009; Keating et al. 2014).

The abundance of low-ionisation metals such as C II and O I is also observed to decrease with redshift over $1.5 < z < 5.5$; however, there have been indications that the volume density of O I systems might be stabilising or even increasing at $z \gtrsim 6$ (Becker et al., 2006). This trend may be linked to an evolution in the UVB and in the physical densities of metal-enriched gas in ways that tend to favour lower ionisation states at higher redshifts, even while the overall metal enrichment is lower. Simulations have supported this view (e.g. Finlator et al. 2015, 2016; Oppenheimer et al. 2009), and predict that C II systems might become equally or more abundant than C IV systems at $z \gtrsim 8$.

Mg II has been used extensively to probe metal-enriched gas at lower redshifts ($z \lesssim 2.5$) (e.g., Weymann et al. 1979; Churchill et al. 1999; Nestor et al. 2005; Lundgren et al. 2009; Weiner et al. 2009; Chen et al. 2010a, Ménard et al. 2011; Kacprzak et al. 2011; Churchill

et al., 2013a,b; Mathes et al. 2017). At higher redshifts, infrared surveys by Matejek & Simcoe (2012) and Chen et al. (2016) have revealed that strong (rest-frame equivalent width $W > 1.0$ Å) absorbers decline with redshift at a rate similar to the global star formation rate, whereas weaker ($0.3 < W < 1.0$ Å) systems show little or no decline over $0.4 < z < 7$. These observations, along with comparative studies of absorber and galaxy properties, have led to the hypothesis that strong Mg II systems trace transient phenomena, such as metal-enriched outflows, that track the global star-formation rate, while weak Mg II systems may be more often associated with inflows, or arise from the fragments of older star-formation-driven winds (see summaries in Kacprzak et al. 2011; Matejek et al. 2013; and Mathes et al. 2017).

The goal of this paper is to provide a sensitive survey for metal lines out to $z = 7$. Towards this aim, we have acquired a deep Very Large Telescope (VLT) X-Shooter spectrum of ULAS J1120+0641 (hereafter J1120; Mortlock et al. 2011) which, at a redshift of $z = 7.0842 \pm 0.0004$, (Venemans et al., 2012), is curently the most distant known QSO. In terms of its UV continuum, it is also one of the most luminous known QSO at $z > 6.5$, making it an excellent target for spectroscopic follow-up, and a powerful probe of metal absorbers beyond redshift six. For C IV and C II these are the first observations of intervening metal lines out to $z = 7$, while for Mg II we probe considerably lower equivalent widths than previous studies.

The remainder of the paper is organised as follows. We introduce our X-Shooter spectrum of J1120 in Section IV.2, and our methodology for identifying and measuring intervening metal absorption lines in Section IV.3, which also outlines the analysis techniques used for extracting number densities, column density distribution functions, and cosmic mass fractions. Our results are presented in Section IV.4, in which the implications for C IV, Mg II and C II are analysed in turn and compared to results at lower redshifts and predictions from numerical simulations. The last part of Section IV.4 presents evidence for partial covering of the QSO line-of-sight by associated absorbers in C IV and N V. A summary of our results is given in Section IV.5. Throughout this paper we assume a flat cosmology with $[\Omega_M, \Omega_k, \Omega_\Lambda, h] = [0.3, 0, 0.7, 0.7]$ and equivalent widths are quoted in the rest frame unless explicitly stated. When quoting uncertainties, we give Bayesian 68% credible intervals unless explicitly stated.

IV.2 Data

We obtained a 30-hour VLT/X-Shooter spectrum of J1120 in observations spanning March 2011 to April 2014 in ESO programmes 286.A-5025(A), 089.A-0814(A), and 093.A-0707(A). The data were reduce using a suite of custom IDL (Interactive Data Language[1]) routines. Individual exposures were flat-fielded and sky-subtracted using the optimal method described by Kelson (2003). Relative flux calibration was applied to the two-dimensional frames using response curves derived from standard stars. A single one-dimensional spectrum using 10 km s^{-1} bins was then optimally extracted from all exposures simultaneously. Telluric correction was performed using SkyCalc atmospheric transmission models (Noll et al., 2014; Smette et al., 2015; Kausch et al., 2015). Absolute flux calibration was performed by scaling the corrected spectrum to match the VLT/FORS2 and GNIRS spectra of J1120 obtained by Mortlock et al. (2011). For the absorption lines analysis, the region redward of the the Lyα forest was normalized using a slowly-varying spline fit. Slit widths of 0"9 were used in the VIS and NIR arms, giving nominal resolutions of $R = 7450$ and 5300, respectively. Inspection of the tel-

[1] http://www.exelisvis.com

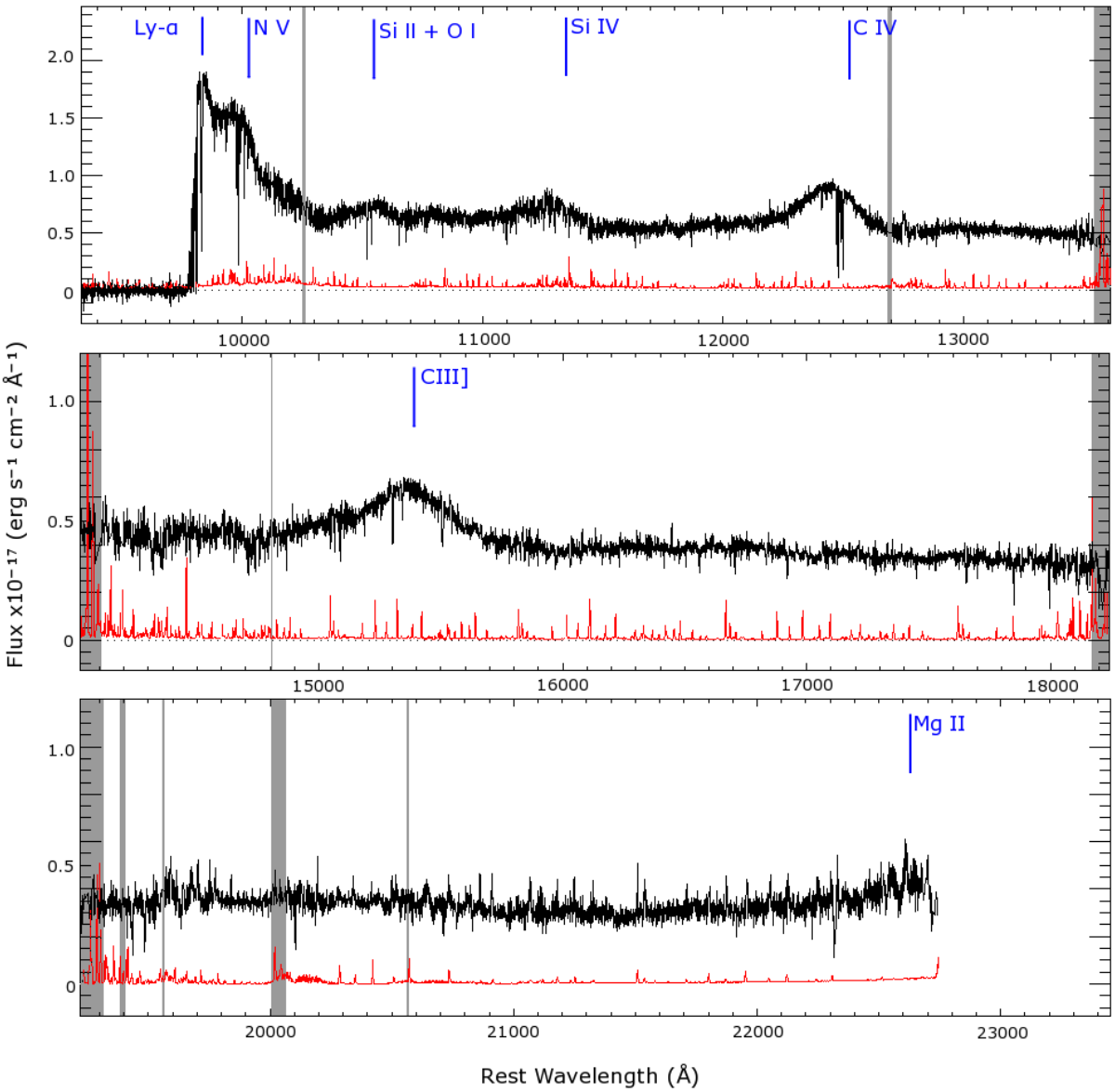

Figure IV.1 – Our 30-hour VLT/X-Shooter spectrum of ULAS J1120+0641. Prominent emission lines are marked assuming a systemic redshift of $z = 7.084$. The spectrum is plotted using 10 km s^{-1} pixels. Areas of overlap between the arms of the spectrograph, as well as regions affected by sky residuals with SNR too poor to allow for the metal search to be conducted are highlighted in grey. To demonstrate the mean SNR, individual pixels affected by skylines have not been plotted. The error spectrum is binned (but not rescaled) in bins of 50 km s^{-1} to avoid crowding the plot with skylines. Some absorption systems can be seen, for instance near the C IV broad emission line.

Table IV.1 – List of the major ion lines included in the search.

Ion	λ/Å
N V	1238.82, 1242.80
O I	1302.16
C II	1334.53
Si IV	1393.76, 1402.77
C IV	1548.20, 1550.77
Si II	1526.71
Al III	1854.71, 1862.79
Fe II	2344.21, 2382.76, 2586.65, 2600.17
Mg II	2796.35, 2803.53
Mg I	2852.96

luric absorption lines, however, indicated that the true mean resolution was somewhat better, $R \simeq 10000$ and 7000, consistent with a typical seeing $\mathrm{FWHM} \simeq 0$"7.

The combined, flux-calibrated spectrum is shown in Figure IV.1. The continuum signal-to-noise ratio spans SNR ~ 10 to 50 per 10 km s^{-1} pixel outside of regions strong affected by the atmosphere. Notably, the residuals from strong sky emission lines tend to exceed the estimates from the formal error array. This appears to be at least partly due to the fact that the projected slit width in the raw frames changes very slightly from one end of the slit to the other. This makes it difficult to model the sky purely as a function of wavelength, and requires an additional fit in the spatial direction. For this we used a variable low-order polynomial; however, non-Poisson residuals still remained for strong lines. We therefore treat regions affected by skylines with caution in our analysis.

IV.3 Search Method

Unlike searches for metal absorption lines at low redshift, where the Lyman-α forest provides a guide as to the redshifts of intervening systems, the onset of nearly complete Lyα absorption at $z \gtrsim 5.5$ means we are forced to rely solely on the metal lines themselves for identification. For this reason we use a modified version of the detection technique used by Becker et al. (2009, 2011), wherein multiple lines are identified with an absorber at a single redshift based on their relative wavelength and optical depth ratios. This process is straightforward for doublets such as C IV $\lambda\lambda 1548, 1551$ and Mg II $\lambda\lambda 2796, 2804$, multiple lines of the same species (as for Fe II), and single lines that commonly arise from the same absorber (e.g., C II $\lambda 1334$ and O I $\lambda 1302$). The final search consisted of the ions listed in Table IV.1. The search window extended from $\lambda = 22750$ Å down to 9838 Å (i.e., the onset of the Lyα forest). Wavelengths affected by strong telluric absorption were masked. The redshift range over which we searched for each ion is shown in Figure IV.4.

We used a two-step process to perform the initial line detection. First, all lines with $\tau \gtrsim 0.3$ were identified visually. A second, automatic line identification procedure was then applied. The automatic algorithm used an inverted Gaussian template, with two free parameters for the depth and width. At regularly spaced intervals in velocity $\Delta v = 20$ km s^{-1}, this template was

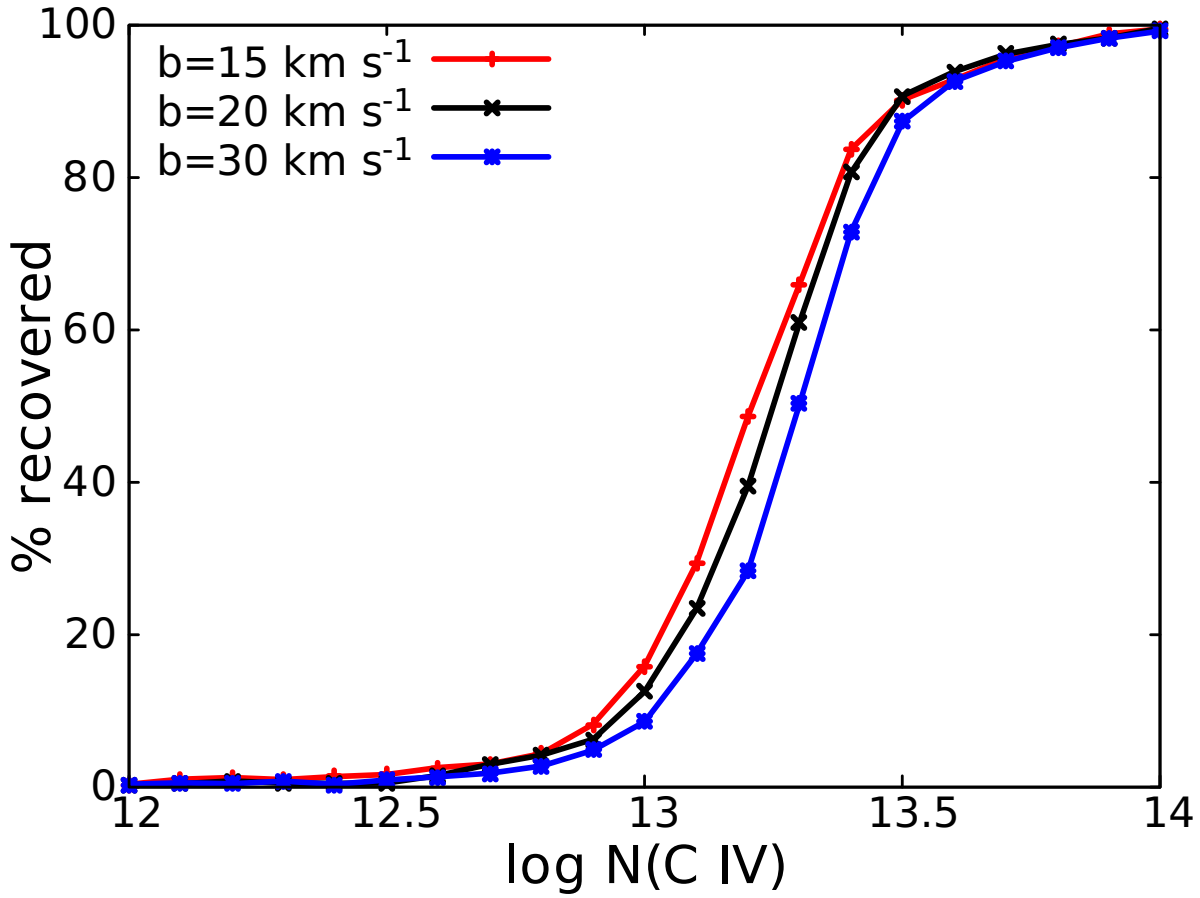

Figure IV.2 – C IV completeness as a function of column density. Each point was obtained by inserting an artificial C IV doublet at a randomly chosen redshift over $5.3 < z < 7.0$ (as plotted in Figure IV.4) 1000 times, and attempting recovery with the automated line detection algorithm described in Section IV.3. The '+', 'x', and '$*$' symbols denote results for $b = 15$, 20, and 30 km s^{-1}, respectively. Completeness estimates over $6.2 < z < 7.0$, over which the fitting is done, are similar.

Table IV.2 – List of detection thresholds for the species of interest. Throughout the paper we make use of the 15 km s^{-1} values for all saturated systems.

Ion	b/km s^{-1}	$\log(N_{\rm min}/{\rm cm}^{-2})$
C IV	15	13.1
	30	13.2
Mg II	15	12.5
	30	12.7
C II	15	13.0
	30	13.2
Si IV	15	12.6
	30	12.7
Mg I	30	12.3
Si II	30	13.2

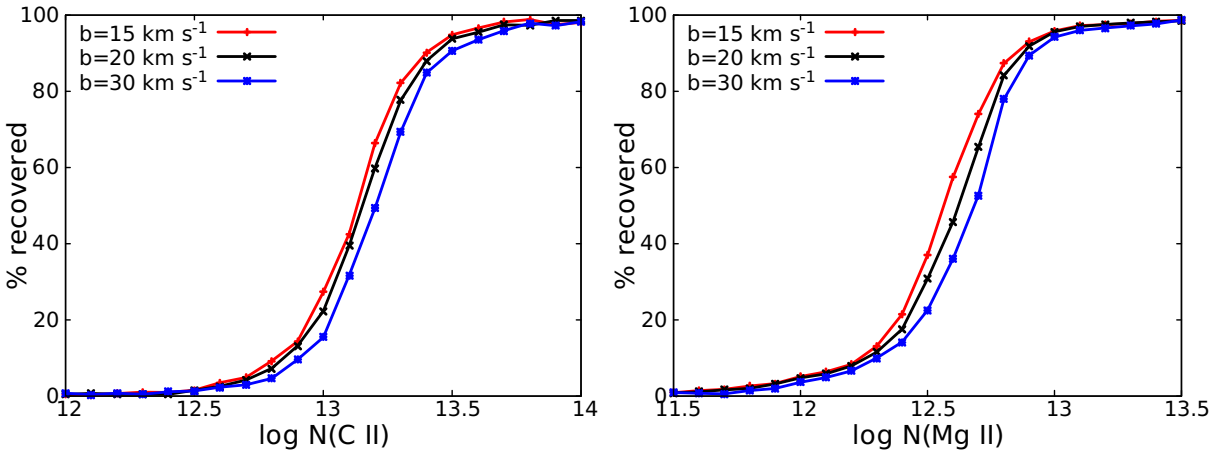

Figure IV.3 – Left: C II completeness as a function of column density. Range searched extends over $6.3 < z < 7.0$. Symbols are as in Figure IV.2. Right: Mg II completeness as a function of column density. Range searched extends over $5.9 < z < 7.0$. Symbols are as in Figure IV.2.

fit to a small region of the spectrum, iteratively rejecting pixels that exceeding a 2σ clipping threshold.

To aid identification, preliminary column densities were derived using the apparent optical depth formula (Savage & Sembach, 1991) in which the optical depth τ is related to the observed flux intensity, F, and the continuum intensity, F_0, as

$$\tau = -\ln\left(F/F_0\right) \,. \tag{IV.1}$$

The ionic column density, N_i, for a pixel of optical depth τ_i is then derived as

$$N_i = \frac{m_e c}{\pi e^2} \frac{\tau_i}{f\lambda_0} \,, \tag{IV.2}$$

where f is the oscillator strength of the relevant ion transition, λ_0 is its wavelength, e the electron charge and m_e is the electron mass. Column densities are quoted in units of cm^{-2} throughout.

The column density thresholds for detection, given in Table IV.2, were chosen so that the number of false detections, following the criteria described below, was either zero or had reached a baseline level driven by bad pixels, where a candidate detection was easily eliminated by visual inspection. These thresholds typically corresponds to column densities where we are $\sim$30 per cent complete. In addition, we required the flux decrement across the selected absorption features to be significant to at least 5 σ based on the noise array.

For positively identified systems, final column density and Doppler b parameters (in km s^{-1}) were obtained by fitting Voigt profiles using the fitting program `vpfit` (Carswell & Webb, 2014). This also allowed us to introduce a variable power-law correction to the continuum for each line. The detected systems are treated as single-component absorption lines for the purposes of measuring the column densities. The few systems which show tentative evidence of multiple components (such as the C IV component of the $z = 6.17$ system) are weak and the existence and strength of additional components is influenced by the data reduction method and initial guesses. These components are also only partially resolved. These uncertainties are

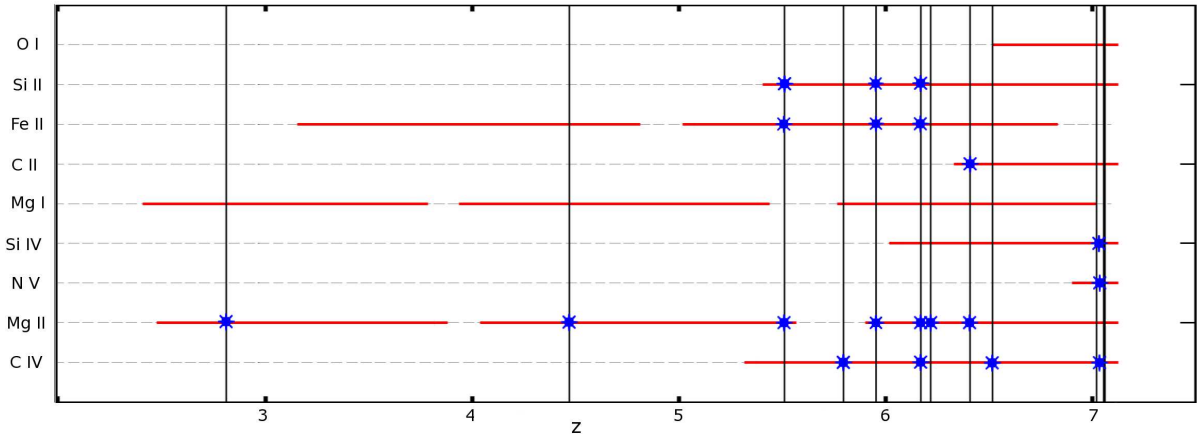

Figure IV.4 – A graphic summary of our survey results. The redshift ranges over which we searched for different ions are shown by horizontal solid lines. The redshifts of detected systems are indicated by vertical solid lines, with detected species marked by blue stars. The emission redshift of the QSO is indicated by a thick vertical line at $z = 7.084$. Nine intervening systems are found, along with three associated systems within 3000 km s^{-1} of the object (shown on the Figure as a single line).

reflected in the errors on log N given in Table IV.3, which encompass the values we would get if we fit multiple components and summed their column densities. In addition, we note that most previous surveys of line-of-sight metals lacked the resolution to notice multiple components in such weak systems. Where required, we computed upper limits for the column density of the undetected ions in identified intervening systems by inserting increasingly strong artificial lines near the corresponding redshift. An offset equal to the Doppler parameter of the detected ions in the system was chosen to mitigate the effect of potential absorption lines just below the detection threshold. The upper limit corresponds to the weakest injected line which would still be detected independently from associated ions, using the same detection criteria near that wavelength.

Detection completeness was evaluated by inserting mock absorbers of each metallic species into the spectrum and attempting to recover them over a range of velocity width parameters and column densities. The redshift ranges probed are the ones shown in Figure IV.4 and vary between ions; the Doppler parameters tested are $b = 15, 20, 30$ km s^{-1} for the most common species C IV, Mg II, C II and Si IV, and $b = 30$ km s^{-1} for less common species. For each combination of b and N, a redshift from the probed range is chosen at random, then an absorption line with those parameters is injected in the spectrum. The search algorithm is then run to attempt to recover the artificial line at 5σ significance based on the error array and at a threshold higher than that of the false positives (see next paragraph). Example results are shown in Figures 2, 3, and 4 for C IV, C II, and Mg II, respectively. We find that we are able to detect > 30 per cent of C IV absorbers of column density $\log(N/\mathrm{cm}^{-2}) = 13.1$ and > 95 per cent at $\log(N/\mathrm{cm}^{-2}) = 13.7$, for $b = 15$ km s^{-1}. Due to the higher oscillator strength of the ion, we are sensitive to 30% of Si IV systems with $\log(N/\mathrm{cm}^{-2}) = 12.7$ and > 95 per cent of those with $\log(N/\mathrm{cm}^{-2}) \geq 13.3$ with Doppler parameter $b = 15$ km s^{-1}. Similarly, we are able to detect $> 30\%$ per cent of Mg II absorbers of column density $\log(N/\mathrm{cm}^{-2}) = 12.5$ and > 95 per cent at $\log(N/\mathrm{cm}^{-2}) = 12.9$ for $b = 15$ km s^{-1}.

Table IV.3 – Intervening systems along the ULAS J1120+0641 line-of-sight. A dash '–' in the column density column indicates that the ion would occur outside of the range probed by our spectrum. Upper limits are given for non-detections. Doppler parameters of less than 15 km s^{-1} indicate that the absorption feature is unresolved with X-Shooter.

z_{abs}	log $N_{C\,IV}/cm^{-2}$	$b_{C\,IV}$	log $N_{Mg\,II}/cm^{-2}$	$b_{Mg\,II}$	log $N_{Mg\,I}/cm^{-2}$	log $N_{C\,II}/cm^{-2}$	log $N_{Fe\,II}/cm^{-2}$	log $N_{Si\,II}/cm^{-2}$
6.51511	$13.25^{a} \pm 0.06$	12 ± 8	< 13.0		< 12.0	< 12.3	< 12.7	< 11.9
6.40671	< 13.1		12.8 ± 0.2	9 ± 6	< 11.6	13.4 ± 0.4	< 12.1	< 12.0
6.21845	< 13.2		12.57 ± 0.07	12 ± 9	< 11.7	–	< 12.4	< 11.9
6.1711	13.67 ± 0.03	57 ± 6	13.4 ± 0.6	19 ± 5	< 11.8	–	12.7 ± 0.2	13.2 ± 0.4
5.9507	< 13.2		12.8 ± 0.1	10 ± 2	< 12.0	–	12.8 ± 0.1	13.2 ± 0.2
5.79539	13.97 ± 0.03	40 ± 2	–	–	< 12.0	–	< 11.5	< 13.1
5.50793	< 12.7		13.37 ± 0.04	49 ± 5	–	–	13.1 ± 0.1	13.5 ± 0.1
4.47260	–	–	12.89 ± 0.04	13 ± 2	< 11.6	–	< 11.7	–
2.80961	–	–	12.81 ± 0.06	12 ± 4	< 11.8	–	–	–

[a] This column density is for the Voigt profile fit to the region in Figure IV.12 that is uncontaminated by skyline residuals in both C IV transitions. Integrating the apparent optical depths over the full λ1548 profile gives $\log N_{C\,IV}/cm^{-2} = 13.46 \pm 0.04$.

To mitigate contamination from false positives, we chose a column density threshold for each species above which detections are considered to be reliable. We determined this threshold by estimating the number of false positives, for a range of ions and column densities, by (i) inserting artificial doublets with incorrect optical depths ratios, based on relative oscillator strengths of the transitions and re-running the detection algorithm to check that no such systems are picked up as valid detections; (ii) inserting artificial doublets with slightly ($\sim 20\%$) incorrect velocity spacing to check the code's sensitivity to spurious interlopers, (iii) searching for doublets with incorrect velocity spacing which should not exist, to check that the code does not pick up chance alignments of noise fluctuations, and (iv) inverting the sign of the Gaussian template to look for spurious lines in emission, and checking that no such features are detected above the chosen column density thresholds. Finally, we ran the code on wavelength ranges visibly devoid of absorbers to check the results were in good agreement with visual inspection.

IV.4 Results

IV.4.1 Overview

In total we detect twelve systems. Nine of these are intervening, and seven of these intervening systems are located at $z > 5.5$. We identify three associated absorbers within 3000 km s^{-1} of the QSO redshift. We do not include these three systems in out main sample; however, two of the absorbers, which appear to be associated with the QSO itself, are analysed further in Section IV.4.6. A summary of the absorber properties is given in Table IV.3.

Plots of the intervening systems can be found in Section IV.6. The two systems at $z < 5.5$ are detected through Mg II only, with tight upper limits on Mg I. No other ions are covered over this redshift range. Meanwhile the seven $z > 5.5$ objects display a wide range of ionic ratios, with five of them displaying Mg II and three of them displaying C IV up to a redshift of $z = 6.515$, currently the highest redshift detection of an intervening metal absorber along a QSO line-of-sight. Fe II and Si II are found in systems located at $z = 5.508$, 5.950, and $z = 6.1714$, while C II is detected with Mg II at $z = 6.406$. Notably, our spectral coverage would allow us to detect Si IV systems located at $z > 6.0$, but none are detected.

Finally, no low-ionisation absorbers are seen in the redshift interval where O I would be detected. We therefore find no apparent overabundance of O I along this line-of-sight, despite indications that such systems might become more common at higher redshifts (Becker et al., 2011). This may be due to scatter between lines-of-sight and a narrow visibility interval ($6.6 < z < 7.0$, corresponding to an absorption path length $\Delta X = 2.0$). On the other hand, we do detect a significant number of Mg II systems at $z > 5.9$, as discussed below.

IV.4.2 Statistics

We compute a range of standard statistics for different metal species. The number density of absorbers is computed alternately per unit redshift, Δz, and per unit absorption path length, ΔX, where

$$X(z) = \int_0^z (1+z')^2 \frac{H_0}{H(z')} \mathrm{d}z' \tag{IV.3}$$

(Bahcall & Peebles, 1969). The column density distribution function (CDDF),

$$f(N) = \frac{\partial^2 n}{\partial N \partial X} \tag{IV.4}$$

can be integrated to obtain the cosmic mass density for a species, usually expressed as a fraction of the critical mass density, $\rho_{\rm crit} = 1.88 \times 10^{-29} h^2$ g cm^{-3}, as

$$\Omega_{\rm ion} = \frac{H_0 m_{\rm ion}}{c \rho_{\rm crit}} \int {\rm d}N N f(N)\,. \tag{IV.5}$$

In practice the mass fraction is computed over a limited range of column densities. We correct for completeness when computing these quantities, as described below.

IV.4.3 C IV

Previous C IV studies have shown a decline in the comoving mass density of C IV between $z \sim 1.5$ and $z \sim 4$ (e.g., D'Odorico et al. 2010, Boksenberg & Sargent 2015), with a possible acceleration of the decline from $z \sim 4.5$ to $z \sim 5.5$ (Becker et al. 2009, Ryan-Weber et al. 2009, Simcoe et al. 2011, D'Odorico et al. 2013). The column density distribution function of absorbers is normally described by a power law whose slope is roughly consistent across $1.5 < z < 5.5$; however, the normalisation of the power law falls by a factor of $\sim$10 with redshift over the same range (D'Odorico et al., 2013). On the modeling side, this has been interpreted as the result of ongoing carbon enrichment in the vicinity of the host galaxies, coupled with a softening of the UVB towards higher redshifts (e.g., Oppenheimer et al. 2009; Oppenheimer et al. 2009; Finlator et al. 2015, 2016).

We use the J1120 line-of-sight to assess whether the observed decline in $\Omega_{\rm C\,IV}$ continues at $z > 6.2$, the highest redshift probed by earlier surveys. Our C IV search above this redshift extends over $6.2 < z < 7.0$, corresponding to $\Delta X = 4.0$. We find only one C IV absorber in this range, at $z = 6.515$. The blue edge of the λ1551 profile is affected by skyline residuals (Figure IV.12). A Voigt profile fit to the the velocity range uncontaminated in both doublet transitions gives $\log N = 13.25 \pm 0.06$. Integrating the optical depths over the full apparent λ1548 profile and applying equation (IV.2), however, gives $\log N = 13.46 \pm 0.06$. In what follows we will general take the lower column density for this system, but we note how our results would change if we adopted the higher value. We also detect two intervening C IV systems at $z < 6.2$. These detections are consistent with previous number density estimates in the literature, but does not provide significant additional constraints since this redshift range has been previously targeted by more extensive surveys.

The constraints which can be obtained from a single detection are naturally weak. Nevertheless, we explore what constraints can be placed on the C IV column density distribution and comoving mass density from our data. Our binned column density results, corrected for completeness, are shown in Figure IV.5. We also show lower-redshift data from D'Odorico et al. (2013), along with power-law fits[2] of the form

$$f(N) = f_0 \left(\frac{N}{N_0} \right)^{-\alpha} . \tag{IV.6}$$

[2]The dashed line in Figure IV.5 for $5.3 < z < 6.2$ is our own fit to the binned data from D'Odorico et al. (2013). We find a slope consistent with their value of $\alpha = 1.44$, but a lower best-fit normalization, $f(N_0) = 7.5 \times 10^{-15}$ cm^2.

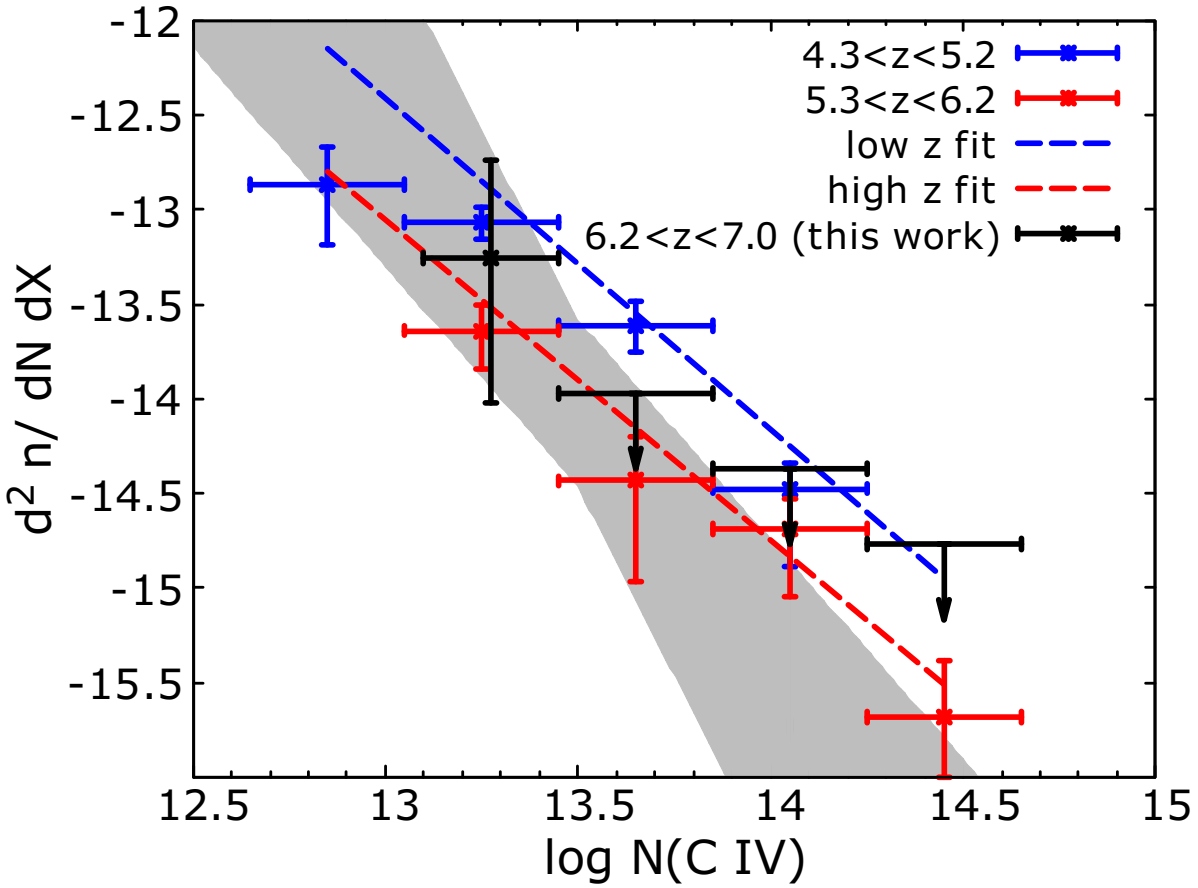

Figure IV.5 – Column density distribution of C IV absorbers at $4.3 < z < 5.3$ (blue, D'Odorico et al. 2013), $5.4 < z < 6.2$ (red, D'Odorico et al. 2013 and J1120 line-of-sight) and $z > 6.2$ (black, this work). Power-law fits are shown as dashed lines. Given our pathlength and completeness, the detection of only a single system in our data, with column density $\log(N_{\rm C\,IV}/{\rm cm}^{-2}) = 13.25$, is consistent with the column density distribution at $z \sim 5.5$, but marginally inconsistent with the $z \sim 4.5$ distribution. The black upper limits correspond to 84 per cent single-sided Poisson uncertainties (Gehrels, 1986). The gray contours show the 68% confidence fit to the column density distribution (see text and Fig. IV.7).

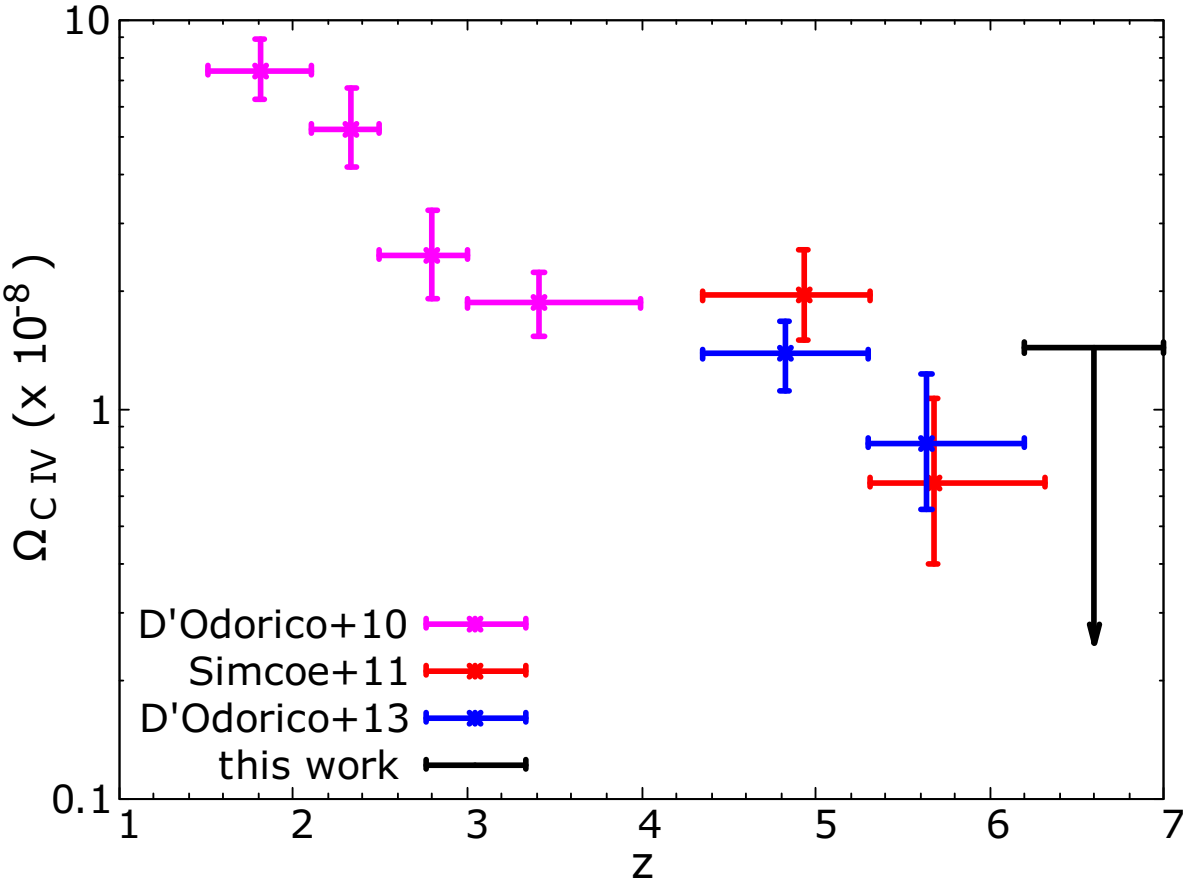

Figure IV.6 – Mass fraction of C IV with redshift, including only strong absorption systems ($13.4 < \log(N_{\rm C\,IV}/{\rm cm}^{-2}) < 15.0$). Our constraints are based on integrating over this column density range after using one detections with $\log(N_{\rm C\,IV}/{\rm cm}^{-2})$ = 13.25 to put constraints the underlying CDDF slope.

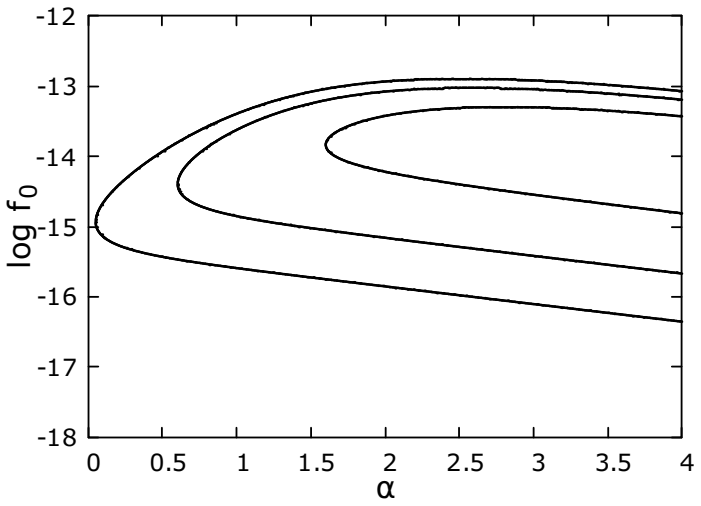

Figure IV.7 – Posterior distribution of the C IV distribution function parameters. Contours correspond to 68%, 95% and 99% credible regions. The fit is made over $6.2 < z < 7.0$ and $\log N > 13.1$, corresponding to our sensitivity threshold. Flat priors on α and $\log N_0$ are adopted (see text).

We choose a fixed pivot value of $\log N_0$ = 13.5. We are roughly consistent with the column distribution of absorbers at $5.3 < z < 6.2$. To obtain constraints on $\Omega_{\rm C\,IV}$, we fit the column density distribution using a maximum-likelihood approach that jointly constrains the amplitude and slope of the column density distribution. We define the likelihood function to be

$$\mathcal{L}(f_0, \alpha) = P_n(n|f_0, \alpha) \times \prod_i P_i\left(N_i|\alpha\right) \ , \tag{IV.7}$$

where P_n is the Poisson probability of observing the total number of systems in our sample, and P_i is the probability of obtaining the the $i^{\rm th}$ column density. All values are corrected for completeness. For each value of α P_i is taken from a distribution where f_0 has been chosen such that the expected mean number of systems is equal to the observed number. Previous works have used a maximum likelihood approach to fit the slope, and then scaled the amplitude of the distribution assuming Poisson statistics (e.g., Matejek & Simcoe 2012).

The advantage of the present approach is that it properly accounts for the degeneracy between α and f_0, which is particularly important for small samples, and does so without binning the data. We verified that our approach recovers appropriate best-fit values and credible intervals using mock samples. We adopt a flat prior on α of $-4 \leq \alpha \leq 0$, which is equivalent to assuming that the distribution has not evolved dramatically from $z \sim 5.5$, for which D'Odorico et al. 2013 find $\alpha \simeq 1.4$. We fit over column densities $\log N_{\rm C\,IV} \geq 13.1$, and use our completeness function for $b = 15\,{\rm km\,s^{-1}}$. The posterior distribution for α and f_0 is shown in Figure IV.7. The marginalised contraints on individual paramters are $\log f_0 = -13.84^{+0.38}_{-0.52}$ and $\alpha > 2.32$ at 68% confidence. We use equation (IV.5) to convert these results into constrains on $\Omega_{\rm C\,IV}$. Integrating over $13.4 \leq \log N_{\rm C\,IV} \leq 15.0$, we find $\log \Omega_{\rm C\,IV} = -8.7^{+1.6}_{-0.5}$. Repeating the analysis using $\log N_{\rm C\,IV} = 13.46$ for the system at $z = 6.515$, we find $\log \Omega_{\rm C\,IV} = -8.5^{+0.7}_{-1.5}$. Adopting the higher column density naturally produces a higher $\Omega_{\rm C\,IV}$, although the upper and lower

limits increase only by a factor of two. Our results are consistent with a continuing decline in $\Omega_{\rm C\,IV}$ with redshift (Figure IV.6), albeit with large errors. The implications of this result are discussed briefly in the next section. We note that models with increasingly negative α become indistinguishable; this is a consequence of the intrinsic degeneracy arising from fitting our data with a power law, and is reflected in the contours in Figure IV.7.

IV.4.4 C II

Our sole C II detection occurs at $z_{\rm abs} = 6.4067$, and is identified via coincidence with Mg II absorption. The column density is $\log N_{\rm C\,II} = 13.4 \pm 0.4$, where our completeness is $\sim$90 per cent. Becker et al. 2011 find an incidence rate of $\mathrm{d}n/\mathrm{d}X \approx 0.25$ at $5.3 < z < 6.4$ using data of comparable quality to our spectrum of J1120 (although the system here is at the lower end of the range of $N_{\rm C\,II}$ in that study).[3] For a non-evolving population we would expect to detect $\sim$1 system over our C II pathlength of $\Delta X = 3.9$. Our data thus presents tentative evidence that the number density of low-ionisation systems remains roughly constant[4] over $5 < z < 7$, which is in turn similar to the number of low-ionisation systems traced by neutral hydrogen absorbers with column densities $\log N_{\rm H\,I} \geq 19.0$ (damped Lyman-α (DLA) and sub-DLA absorbers) over $3 < z < 5$.

Although the statistical claims that can be made from a single line-of-sight are naturally limited, our data point to an evolution in carbon over $6 \lesssim z \lesssim 7$ where the mass density of highly ionised C IV declines with redshift while the number density of low-ionisation absorbers traced by C II remains roughly constant. This is broadly consistent with recent numerical models of chemical enrichment by star formation-driven galactic winds where the metals occupy increasingly higher densities and are exposed to an increasingly softer, spatially fluctuating ionizing background towards higher redshift (e.g., Oppenheimer et al. 2009; Finlator et al. 2015, 2016). Significant numerical challenges remain in modeling these absorbers (e.g., Keating et al. 2016; Chen et al. 2016); however, the data presented here should provide additional leverage for testing numerical models in terms of their redshift evolution.

IV.4.5 Mg II

Our deep X-Shooter spectrum is the first to be highly sensitive to very weak ($W < 0.3$) Mg II systems out to $z = 7$. We searched for lines over $2.5 < z < 7.0$ with gaps around $3.9 < z < 4.1$ and $5.6 < z < 5.9$ due to telluric absorption (see Fig. 3). We detect five Mg II systems at $z > 5.5$ (four at $z > 5.9$), all of which have $W < 0.5$ Å, and three of which show additional low-ionisation ions (see Table IV.3).

The equivalent width distribution of Mg II absorbers has been shown (e.g., Nestor et al.

[3] We note that, unlike absorption doublets, singlet species such as C II cannot be identified on their own. Without the Lyman-α forest to flag potential low-ionisation absorbers via their H I absorption, these ions must be identified via coincidence with other metal lines. In this sense, we caution that the detection method for C II is not consistent across surveys. Becker et al. 2011 lacked the near-infrared coverage to detect Mg II, and searched for C II based on coincidence with Si II and O I, while our C II system was detected concurrently with Mg II. Care may therefore need to be taken when evaluating trends in C II between different studies.

[4] Although see Becker et al. 2011, who find that the incidence rate of low-ionization systems may increase at $z \gtrsim 5.7$.

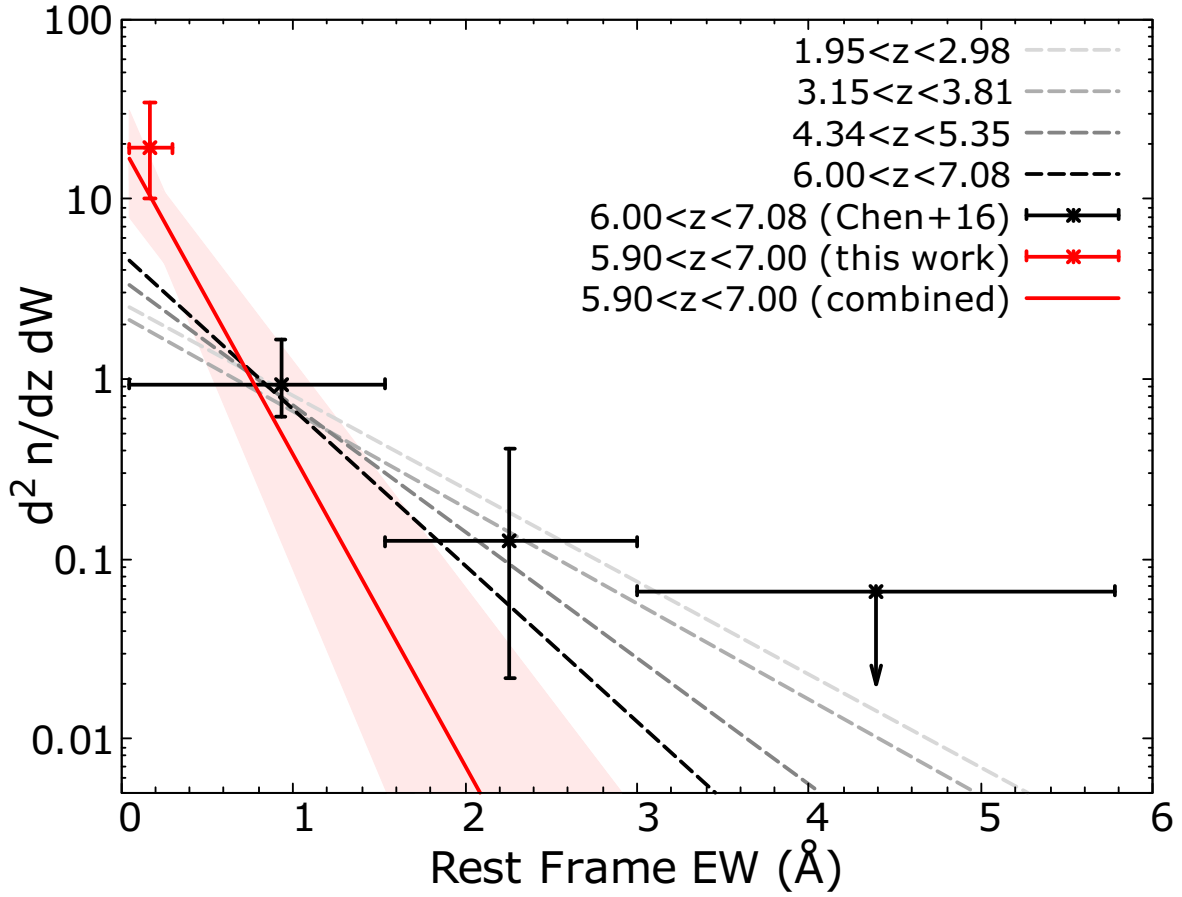

Figure IV.8 – Equivalent width distribution of Mg II absorbers. Data points are from Chen et al. (2016) and this work. Error bars assume Poisson statistics. The dashed lines show the fits to the distribution at different redshifts from Chen et al. (2016). The solid line is our fit to the distribution over $5.9 < z < 7.0$ using the combined datasets. The shaded region is the 68 per cent credible region in the fit. As discussed in the text, a single power law may not provide a sufficient description of the equivalent width distribution over the full range in W.

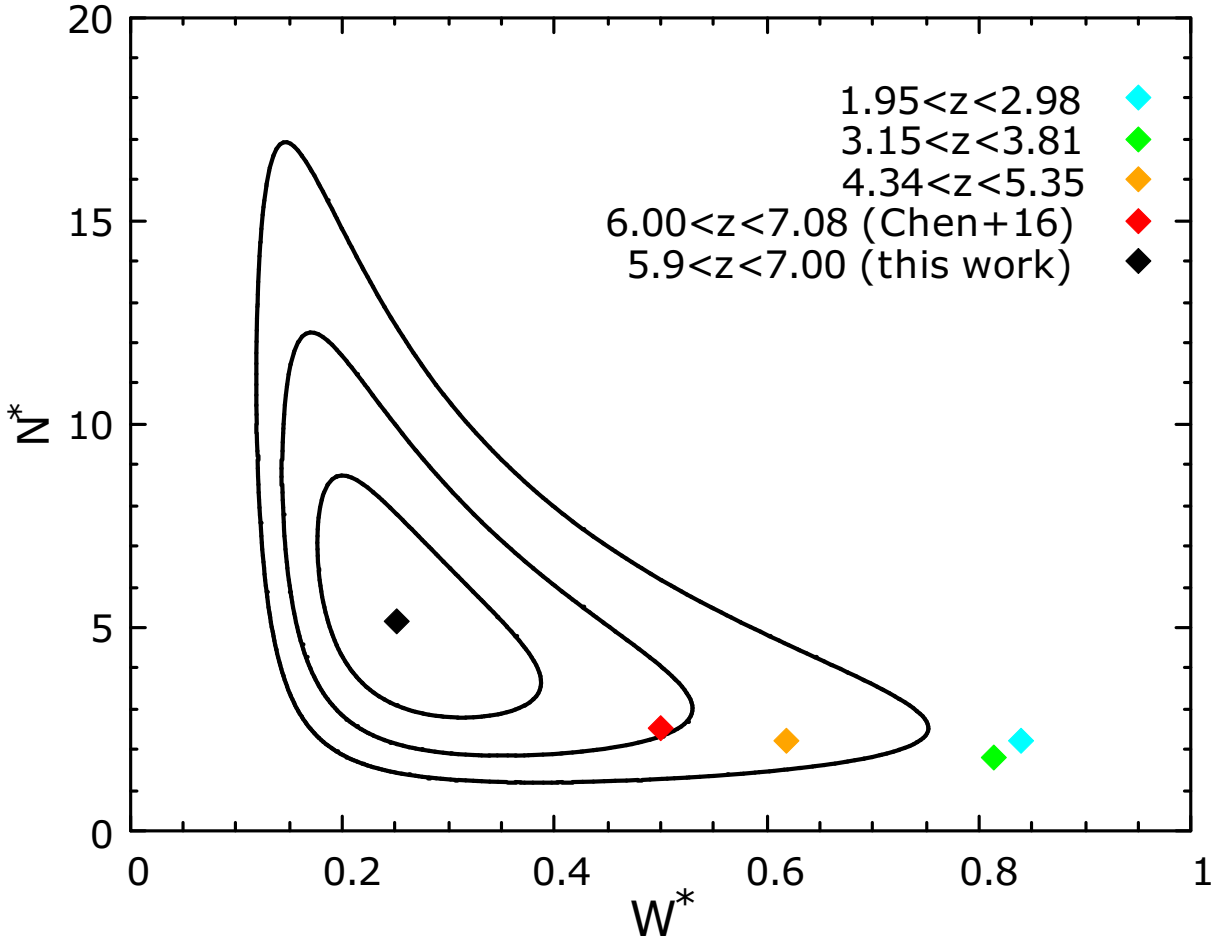

Figure IV.9 – Posterior distribution of the Mg II distribution function parameters. The best fitting parameters are indicated by a black diamond, with black contours corresponding to 68%, 95% and 99% credible regions. The fit is made over $5.9 < z < 7.0$ (see text). Best fit parameters obtained by Chen et al. (2016) at $\overline{z} = 2.52, 3.46, 4.80, 6.28$ are shown as coloured diamonds. Our measurement is in 2σ tension with previous work at the same redshift.

Table IV.4 – List of equivalent widths for Mg II systems detected along the line-of-sight. Errors are measured from the error array.

z	$W/\text{Å}$
6.40671	0.094 ± 0.022
6.21845	0.139 ± 0.029
6.1711	0.258 ± 0.057
5.9507	0.188 ± 0.024
5.50793	0.455 ± 0.056
4.47260	0.276 ± 0.012
2.80961	0.246 ± 0.020

2005) to be well described by an exponential function of the form

$$\frac{\mathrm{d}^2 n}{\mathrm{d}z\,\mathrm{d}W} = \frac{N^*}{W^*} e^{-W/W^*} , \tag{IV.8}$$

at least for $W > 0.3$ Å, a point we will return to below. The scale factor W^* peaks at $z \sim 2.5$. Using the best-fit parameters from the highest redshift bin of Chen et al. (2016) ($W^* = 0.50$ Å, $N^* = 2.51$ over $6.00 < z < 7.08$; see Figure IV.8), the expected number of systems along the J1120 line-of-sight with $W > 1$ Å over $5.9 < z < 7.0$ is ∼0.4, consistent with our non-detection of strong systems.[5] For the same fit, however, we would expect to detect only ∼0.6 systems with $W < 0.3$ Å, given our completeness, whereas we detect four (all with $W > 0.9$ Å; see Table IV.4). Binned values of $\mathrm{d}^2 n/\mathrm{d}z\mathrm{d}W$ from this work and Chen et al. (2016) are plotted in Figure IV.8. Our binned value is estimated as $\mathrm{d}^2 n/\mathrm{d}z\mathrm{d}w \simeq \left(\sum C_i^{-1}\right) / \left(\Delta z \Delta W\right) = 19.1$, where C_i is our completeness at the column density of the i^{th} absorber, $\Delta z = 1.1$, and $\Delta W = 0.25$, where our bin spans 0.05 Å $< W < 0.3$ Å. The error bars are Poisson.

We re-evaluate the distribution of Mg II systems at $z > 5.9$ by combining our data with those of Chen et al. (2016) excluding their J1120 line-of-sight. The Chen et al. (2016) sample is less sensitive but contains more lines of sight at $z > 6$, and therefore better constrains the the high-W end of the distribution. We use a maximum-likelihood, full Bayesian approach to constrain W^* and N^* similar to the one described in Section IV.4.3. Here, the combined likelihood function is given by

$$\mathcal{L}(W^*, N^*) = P_n(n|W^*, N^*) \times \prod_i P_i\left(W_i|W^*\right) , \tag{IV.9}$$

where P_n is the Poisson probability of detecting the total number of lines in our sample, and P_i is the probability of obtaining the i^{th} equivalent width. We use a redshift path-weighted mean completeness function that combines our J1120 data with the remainder of the sightlines from Chen et al. (2016). The four Mg II systems in our sample are combined with seven from Chen et al. (2016) for a total of $n = 11$. The fit is then performed over 0.05 Å $< W < 5.0$ Å.

Our two-dimensional likelihood values are shown in Figure IV.9. We find best-fitting values of $W^* = 0.25^{+0.09}_{-0.06}$ Å, $N^* = 5.128^{+1.75}_{-1.78}$, where the errors are 68 per cent marginalized credible regions. The combined best-fit values of Chen et al. (2016), which used only stronger systems, are excluded at the ∼93 per cent level (Figure IV.9). It is not clear, moreover, that a

[5] Chen et al. (2016) also found no strong Mg II systems towards J1120.

single exponential provides a good fit over the full range of equivalent width. With our best-fit parameters, for example, we would expect to detect ∼1.5 systems with $W < 0.3$ Å, given our completeness, whereas we detect four, which would have a ∼5 per cent probability of occurring by chance for purely Poisson statistics. Our results were confirmed by a recent, larger study of quasar lines of sight at by Codoreanu et al. (2017), who find that the number of weak Mg II systems becomes inconsistent with a power law extrapolation from strong systems at $z > 4.8$ – although this might also be the case at lower redshifts. Indeed, Nestor et al. (2005), using observations by Churchill et al. (1999), first pointed out that the equivalent width distribution of Mg II is more complicated than a single power law. They fit a double exponential at $0.4 < z < 1.4$, finding an upturn in the number density of systems below $W = 0.3$Å. More recently, Mathes et al. (2017) used a Schechter function to fit the equivalent width distribution at $z < 2.6$, finding an exponential cut-off near $W \sim 2$ Å.

While the distribution function of Mg II has not been probed in the weak regime at intermediate redshifts ($2.5 \lesssim z \lesssim 6$), the apparently high number of weak Mg II systems we detect plausibly reflects complexity in the shape of the equivalent width distribution similar to what is seen at lower redshifts ($z \lesssim 2.5$). In terms of their physical properties, weak Mg II systems may not have the same origin at all redshifts. Even so, it is possible that these weak systems at $z \sim$ 6-7 are associated with accreting gas and/or the cooling remnants of previous metal-enriched outflows, as has been suggested for weak systems at $z \lesssim 2.5$ (see discussion in Mathes et al. 2017).

IV.4.6 Associated Absorbers

In addition to intervening absorbers along the line-of-sight to J1120, we analysed absorbers close to the redshift of the QSO. We find three such systems, located at -2530, -1100 and -920 km s^{-1} blueward of the systemic redshift. The strongest system, at -1100 km s^{-1}, was identified in C IV and N V in the discovery spectrum by Mortlock et al. (2011). The two remaining systems, as well as the Si IV in the strongest system, are newly identified here.
The system at $z = 7.01652$ (2530 kms^{-1} from the QSO's systemic redshift) contains weak C IV absorption and can be seen in Figure IV.21. The two highest-redshift systems, at $z = 7.05540$ and $z = 7.06001$, consist of C IV and N V, as well as Si IV absorption in the former.[6]

The associated systems at $z \simeq 7.06$ display unusual absorption profiles, as the apparent optical depths of the C IV and N V doublets are in ratios $\tau_{1548}/\tau_{1550} = 1.32$, $\tau_{1238}/\tau_{1242} = 1.21$, respectively, for the $z = 7.060$ system and $\tau_{1548}/\tau_{1550} = 1.05$, $\tau_{1238}/\tau_{1242} = 1.15$ for the $z = 7.055$ system. While saturation can drive the equivalent width ratios below the canonical value of 2:1 expected for optically thin lines, for the $z = 7.055$ system the large residual flux (~ 40 per cent) makes it unlikely that saturation is the only effect (see Figures IV.10 and IV.11). Two plausible explanations of the peculiar ratio of the doublets are (i) the intervening system contains a column density of C IV and N V sufficient for saturated absorption, but covers only a fraction of the continuum source – *partial covering*, or (ii) the absorption feature is composed of multiple unresolved components, each individually saturated and located close enough in velocity space to appear blended.

The `vpfit` program was adjusted to test the relative goodness of fit provided by these two

[6]The fact that these systems occur blueward of the QSO redshift yet the C IV lines fall on the red side of the C IV emission line (Figure IV.1) reflects the extreme blueshift of this object's C IV emission line, noted by Mortlock et al. (2011).

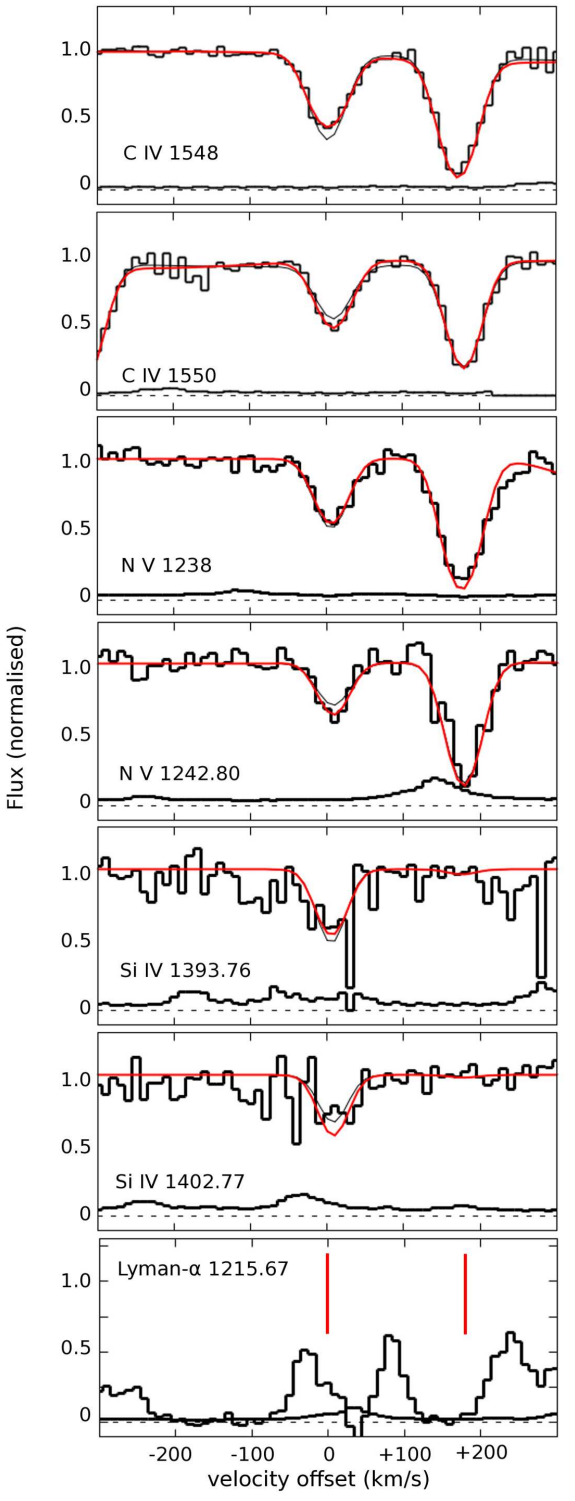

Figure IV.10 – Associated absorption systems at $z = 7.055$ and $z = 7.060$. The plot is centered at $z = 7.055$. Thick continuous lines show single-component fits using partial covering (Table IV.6). Thin lines show fits without partial covering (Table IV.5). The locations of Lyman-α for these components are indicated with tick marks in the bottom panel.

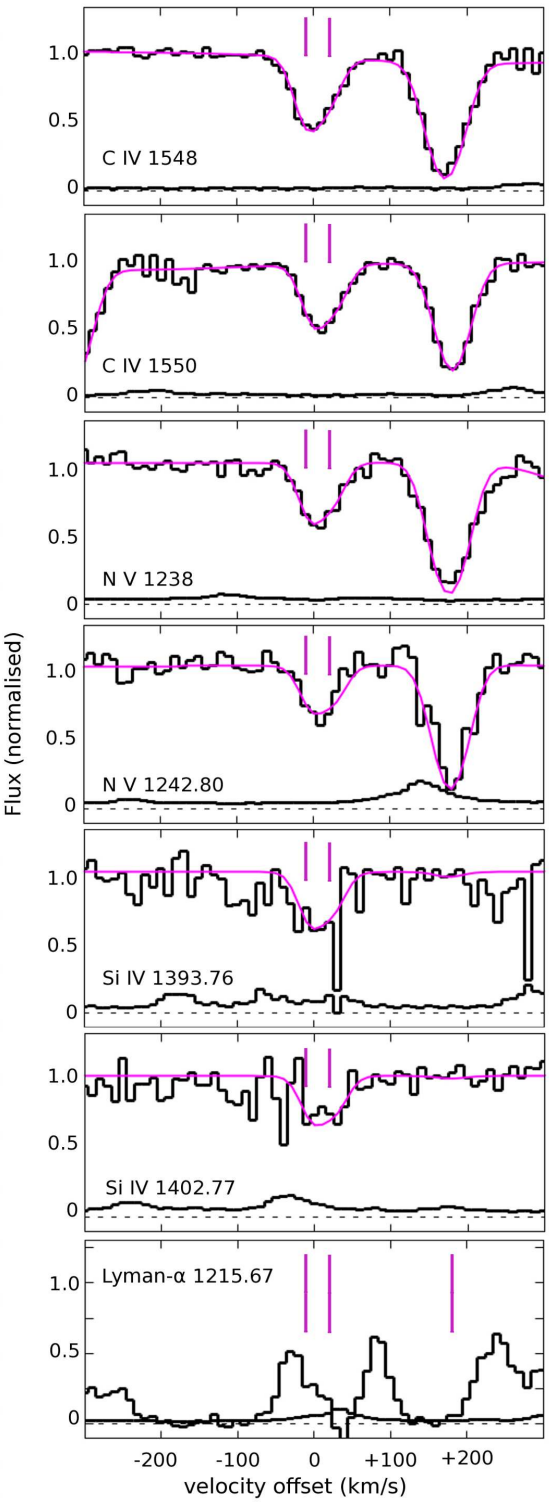

Figure IV.11 – Same as Figure IV.10, but here the thick continuous lines show fits using multiple, unresolved components for the system at $z = 7.055$ (Table IV.7). The velocities of the components are indicated with vertical tick marks.

Table IV.5 – Best-fit parameters for the fits to associated systems using single components and no partial covering.

$z_{\rm abs}$	log $N_{\rm C\,IV}/{\rm cm}^{-2}$	log $N_{\rm N\,V}/{\rm cm}^{-2}$	log $N_{\rm Si\,II}/{\rm cm}^{-2}$	b
7.05541 ± 0.00002	13.88 ± 0.02	13.87 ± 0.03	13.34 ± 0.09	21.8 ± 1.4
7.060000 ± 0.000013	14.44 ± 0.04	14.7 ± 0.09	11.9 ± 0.5	19.3 ± 1.0

$\chi^2/N_{\rm dof} = 2.950$

$\chi^2_{\rm no\ Si\ IV}/N_{\rm dof} = 2.115$

Table IV.6 – Best-fit parameters for the fits to associated systems using single components and allowing for partial coverage of the continuum.

$z_{\rm abs}$	$F_{\rm cover}$	log $N_{\rm C\,IV}/{\rm cm}^{-2}$	log $N_{\rm N\,V}/{\rm cm}^{-2}$	log $N_{\rm Si\,II}/{\rm cm}^{-2}$	b
7.05541 ± 0.00002	42 ± 2 per cent	14.6 ± 0.2	14.37 ± 0.11	14.2 ± 0.3	17.9 ± 1.8
7.060000 ± 0.000013	10 ± 5 per cent	14.44 ± 0.04	14.82 ± 0.12	11.9 ± 0.5	18.8 ± 1.1

$\chi^2/N_{\rm dof} = 2.641$

$\chi^2_{\rm no\ Si\ IV}/N_{\rm dof} = 1.736$

Table IV.7 – Best-fit parameters for the fits to associated systems using single components and no partial covering.

$z_{\rm abs}$	log $N_{\rm C\,IV}/{\rm cm}^{-2}$	log $N_{\rm N\,V}/{\rm cm}^{-2}$	log $N_{\rm Si\,II}/{\rm cm}^{-2}$	b
7.05514 ± 0.00004	14.2 ± 0.3	13.94 ± 0.15	13.6 ± 0.5	5.8 ± 1.8
7.05596 ± 0.00006	14.3 ± 1.4	14.1 ± 1.0	14.5 ± 1.6	2.7 ± 1.6
7.060002 ± 0.000013	14.45 ± 0.04	14.84 ± 0.12	11.9 ± 0.5	18.7 ± 1.1

$\chi^2/N_{\rm dof} = 2.656$

$\chi^2_{\rm no\ Si\ IV}/N_{\rm dof} = 1.772$

possibilities. To test partial covering, an additional variable representing the covering fraction of the continuum was added to the fit. The C IV, N V and Si IV lines were fit simultaneously, with the covering fraction, redshift and Doppler parameter of each component constrained to be the same for all ions. Multiple narrow absorbers, on the other hand, were fit by modifying the initial conditions of the fit to contain two narrow absorbers for each absorption line. The starting values of the Doppler parameters were initially forced to be $b < 7$ km s^{-1}. After letting the fit converge this condition was relaxed and the fit was re-run. Component redshifts and Doppler parameters were once again tied between ions. In both these fitting techniques, we introduced an extra 'slope' parameter over each fitting window to allow for adjustments to the continuum normalisation.

Narrow absorbers and partial covering provide comparably good fits to the data, both improving upon naive single-component fits (Table IV.5). The best-fit partial covering fraction for the $z = 7.055$ system is 42 ± 2 per cent (see Table IV.6), with a χ^2 per degree of freedom integrated over all components of 2.641. In both this fit and the alternative, the χ^2 is driven primarily by the Si IV doublet, which may suggest that the noise in the Si IV line exceeds the estimate in the error array (see Figure IV.10). Omitting Si IV, the reduced χ^2 is 1.736. Multiple unresolved absorbers provide a similarly good fit to the data, but the column density of one of the components is highly unconstrained for all ions (Table IV.7). The χ^2 per degree of freedom of the fit is 2.656, dominated again by the Si IV doublet, or 1.772 when omitting Si IV .

The χ^2 results do not allow one model to be definitely preferred over the other. The multiple component model, however, may be less viable on physical grounds. Using $b = \sqrt{2kT/m}$, a b-parameter of 2.7 kms^{-1} (4.3 kms^{-1} at the 1σ upper limit) would set an upper limit on the temperature of the C IV gas of $T \lesssim 5000$K ($T \lesssim 12000$K), which is potentially problematic if the gas is photo-ionised by the QSO. Partial covering in associated narrow QSO absorbers, on the other hand, is a well-documented phenomenon (D'Odorico et al. 2004; Misawa et al. 2007; Wu et al. 2010; Simon et al. 2012).

The partial covering hypothesis, if correct, could have significant implications for the proximity zone of J1120. As noted by Simcoe et al. (2012), Lyman-α at the redshift of these metal absorbers is unsaturated – our spectrum confirms this. Lack of saturation would normally indicate that the H I column density is too low to be optically thick to ionizing photons ($\log N_{\rm H\,I} < 17.2$). If partial covering is a factor, however, then optically thick H I may indeed be present, but suppressing only part of the QSO continuum. This is more likely for the component at $z = 7.055$, which contains Si IV and is probably less highly ionized than the component at $z = 7.060$. Even a partial suppression of the ionizing continuum could contribute to the apparent shortness of the proximity zone noted by Mortlock et al. (2011). While it is difficult to know whether this scenario is correct for J1120, it may be of interest as data for further QSOs at $z > 7$ are obtained.

IV.5 Summary

We have used a deep (30h) X-Shooter spectrum of the $z = 7.084$ QSO ULAS J1120+0641 to probe absorption by multiple metal species up to the highest redshifts to date. We find seven intervening systems in the range $5.5 < z < 7.0$ and three associated systems. The intervening systems span a wide range of ionic compositions and velocity profiles. Our main results are:

1. We detect a single C IV system at $z > 6.2$, which is a relatively weak absorber (log

$N_{C\,IV}$ = 13.25 ± 0.06) at $z = 6.51$. Using a maximum likelihood method to set limits on the column density distribution, we demonstrate that the inferred comoving C IV mass density at $z > 6.2$ is consistent with a continuous decline over $4 < z < 7$, though non-evolution from $z \sim 5.5$ cannot be ruled out.

2. We find one C II absorber over $6.3 < z < 7.0$, consistent with the incidence rate of low-ionization absorbers at $z \sim 6$. A decline in C IV with redshift and a relatively flat evolution in C II would be consistent with models that combine lower overall enrichment and a softer ionising background towards higher redshifts.

3. We identify four weak ($W < 0.3$ Å) Mg II systems, exceeding predictions based on an extrapolation of a power law fit to the incidence rate of stronger systems at these redshifts (Chen et al., 2016). This is reminiscent of a similar enhancement in the number density of weak systems at $z < 2.5$ (e.g., Nestor et al. 2005), which are potentially associated with inflows and/or cooling fragments of metal-enriched outflows (e.g., Mathes et al. 2017).

4. We also investigate N V, C IV, and Si IV systems associated with the QSO itself. One system located ∼1100 km s^{-1} blueward of the QSO shows peculiar absorption profiles in the C IV and N V doublets in terms of the relative strengths of the doublet lines. Two explanations, partial covering of the continuum source and multiple unresolved components, are tested and found to explain this effect comparably well. Multiple narrow components provide a reasonable fit; however, we argue that this scenario is physically unlikely as it would require photoionised gas within +1000km s^{-1} of the QSO to have a temperature $T \lesssim 5000$ K ($T \lesssim 12000$ K using the upper 1σ bound on b). Alternatively, a single-component absorber with a covering fraction of ∼40 per cent would produce a similar line profile. In this scenario, a partially covered hydrogen Lyman limit system could also be present even though Lyman-α at the redshift of the metal absorber is not saturated. Such a scenario could potentially help explain the apparent shortness of J1120's proximity zone.

It is worth emphasizing that these results are based on only one line-of-sight. We have often estimated uncertainties using Poisson statistics, which may under-estimate the scatter between lines of sight if metal absorbers are significantly clustered at these redshifts. In addition to large-scale density variations, clustering due to fluctuations in the ionising background could also play a role. The recent discovery at $z \sim 5.5$ of a contiguous ~ 110 comoving Mpc trough of opaque Lyman-α absorption by Becker et al. (2015b) illustrates the fact that ionisation conditions in diffuse gas can be correlated over large distances at these redshifts. Nevertheless, while the information which can be gained from a single line-of-sight is limited, it provides a glimpse into the circum-galactic media of some of the earliest galaxies, and therefore into key mechanisms governing galaxy formation. Metal lines are powerful tools for studying the high redshift universe, and future studies should shed further light on the trends hinted at here.

IV.6 Additional figures: Spectra of intervening systems

Here we plot the X-Shooter spectrum of J1120 at the location of the detected absorbers. The range shown covers $\Delta v = \pm 400$ km s^{-1}; the pixel size is 10 km s^{-1}. Shaded regions highlight the detected lines.

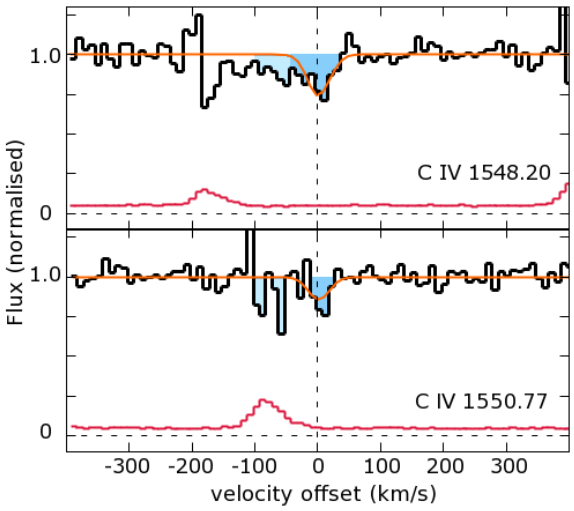

Figure IV.12 – Intervening system at $z = 6.51511$. The thick solid histogram shows the flux. The thin histogram is the error array. Ion identifications are printed at the bottom right of each panel. The orange line shows the best fit Voigt profile returned by `vpfit` (see text) The blue wing of the λ1550.77 line is affected by skyline residuals. The detection of absorption at $\Delta v < -20$ km s^{-1} (shaded) is therefore tentative for this system.

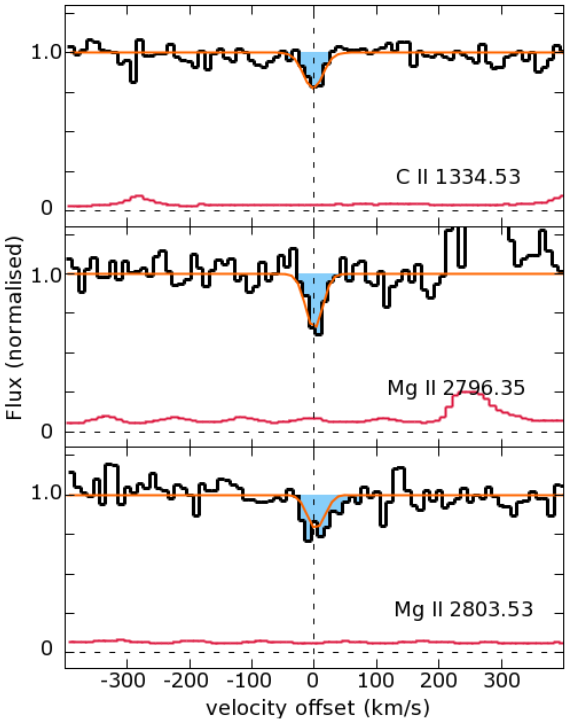

Figure IV.13 – Intervening system at $z = 6.40671$. Lines are as in Figure IV.12.

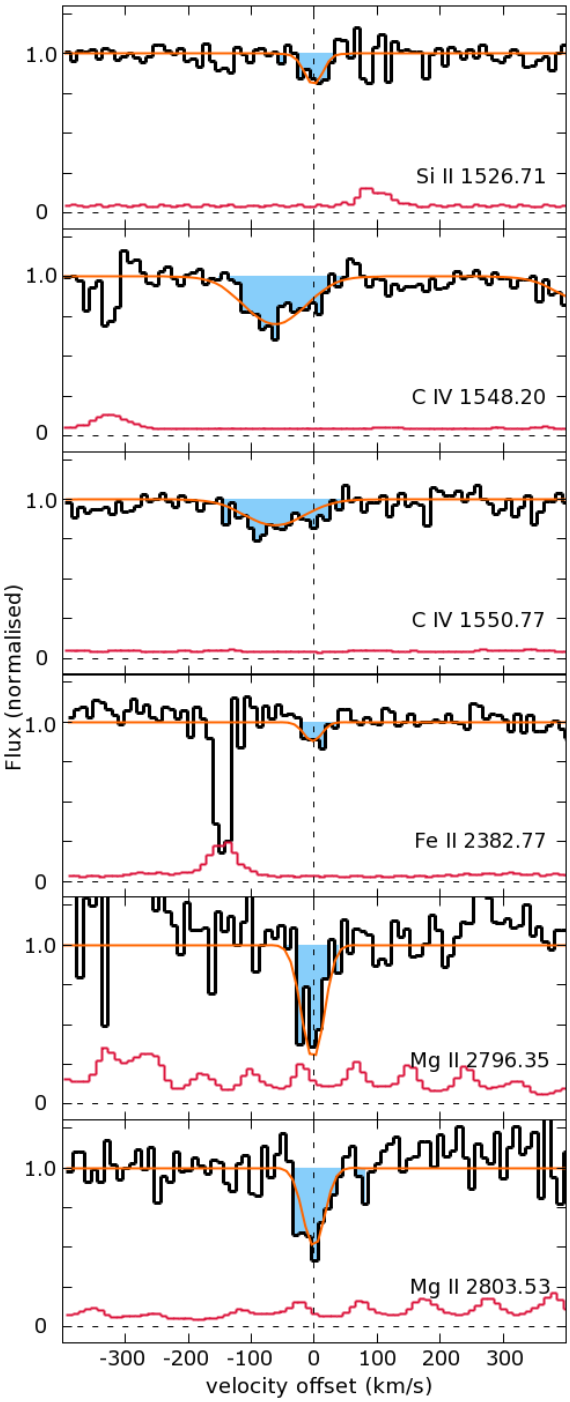

Figure IV.14 – Intervening system at $z = 6.17110$. Lines are as in Figure IV.12.

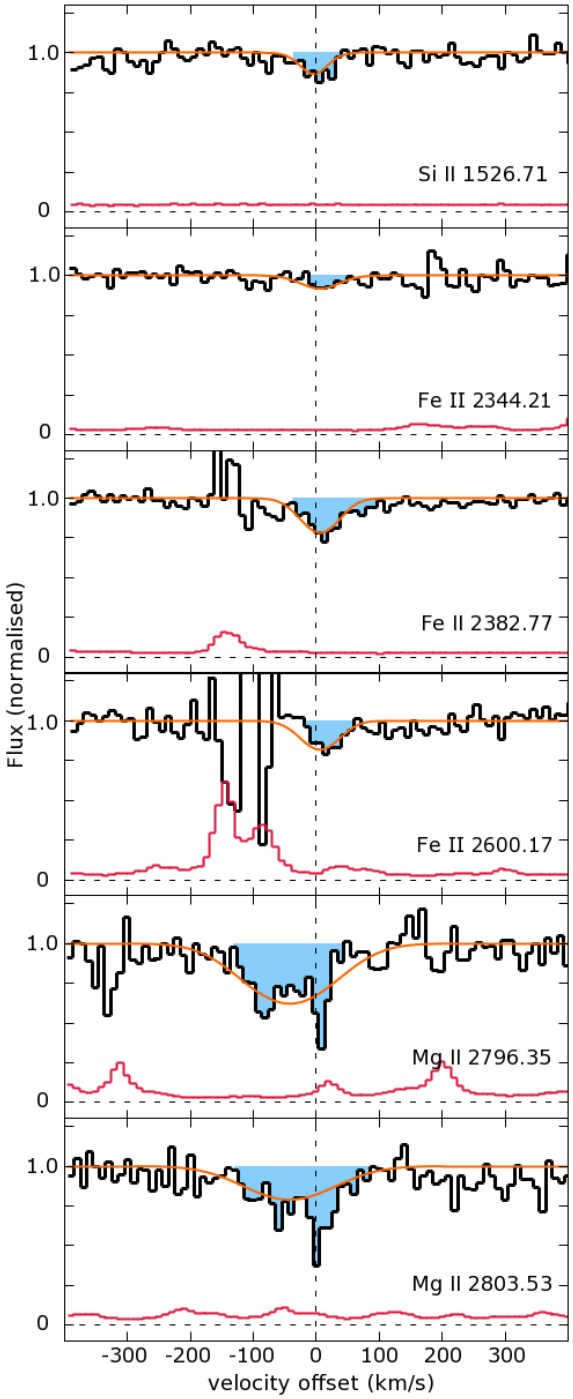

Figure IV.15 – Intervening system at $z = 5.9507$. Lines are as in Figure IV.12.

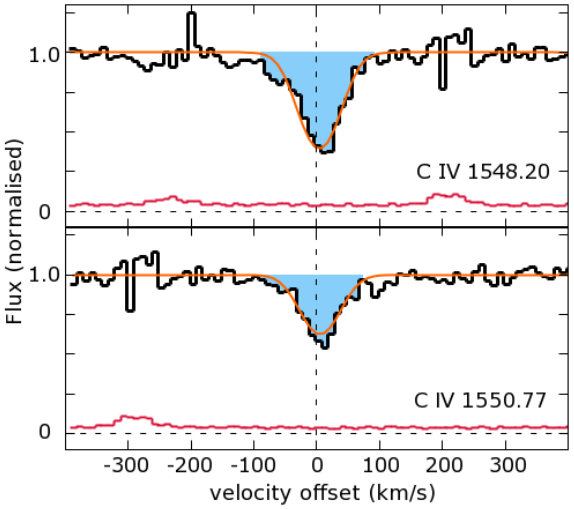

Figure IV.16 – Intervening system at $z = 5.79539$. Lines are as in Figure IV.12.

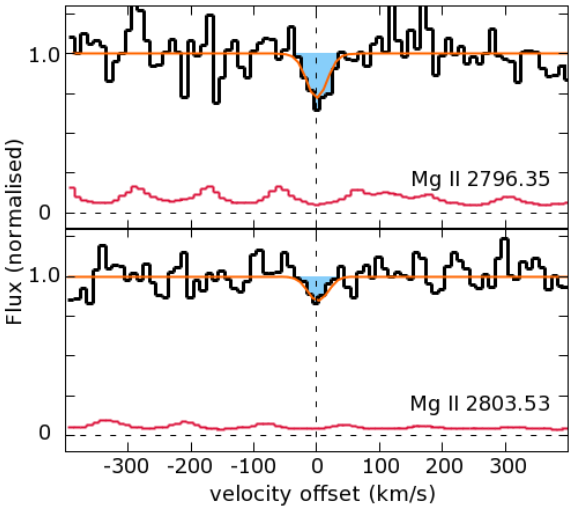

Figure IV.17 – Intervening system at $z = 6.21845$. Lines are as in Figure IV.12.

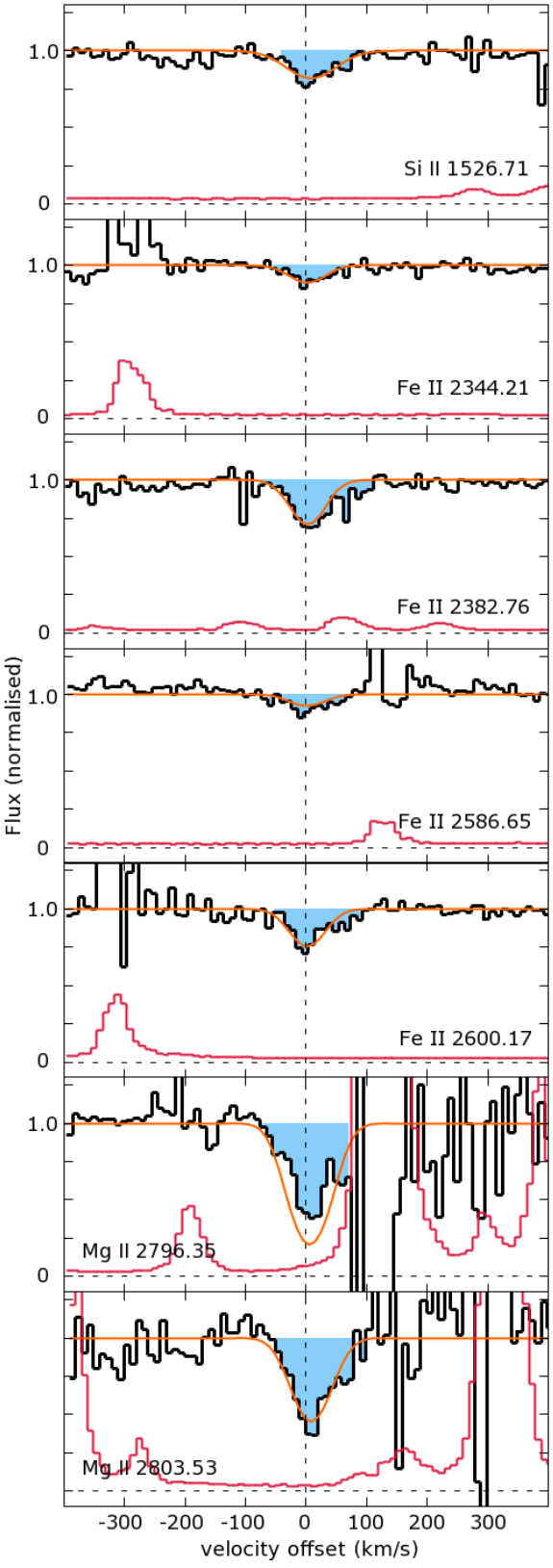

Figure IV.18 – Intervening system at $z = 5.50793$. Lines are as in Figure IV.12. Note that both transitions of Mg II are strongly affected by skyline residuals at $\Delta v \gtrsim 70$ km s^{-1}. In this case, the column density of Mg II is measured using the 2803.53 Å transition alone.

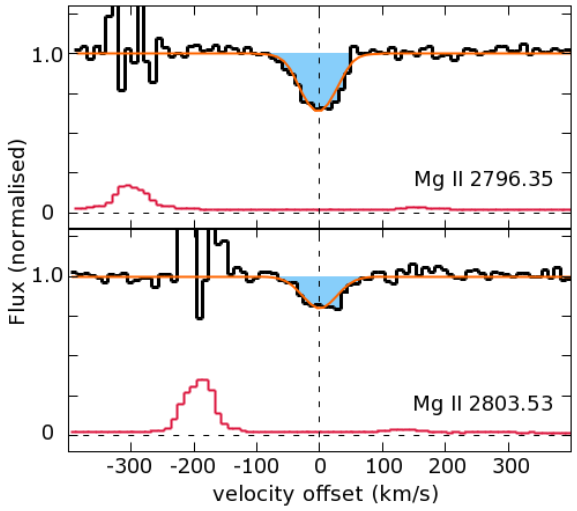

Figure IV.19 – Intervening system at $z = 4.47260$. Lines are as in Figure IV.12.

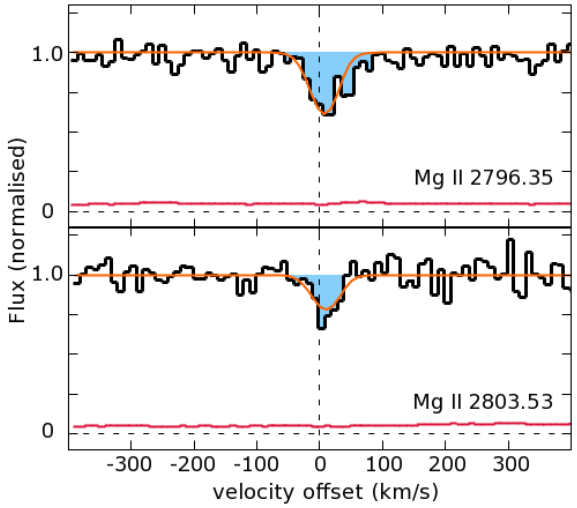

Figure IV.20 – Intervening system at $z = 2.80961$. Lines are as in Figure IV.12.

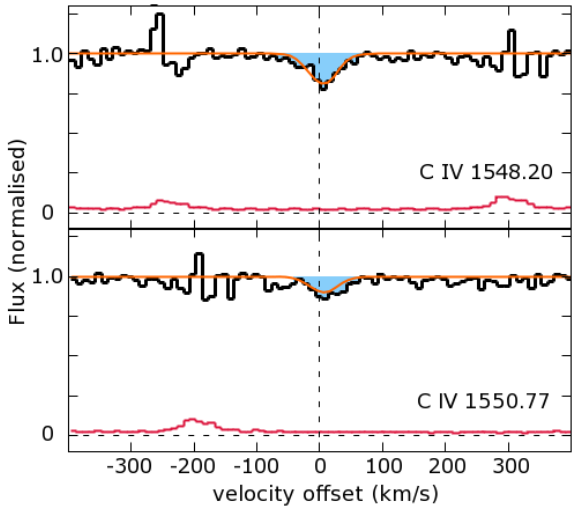

Figure IV.21 – Intervening or possibly associated system at $z = 7.01652$. Lines are as in Figure IV.12.

V

New constraints on Lyman-α opacity with a sample of 92 quasars at $z > 5.7$

Let probability and sample size do the heavy lifting.

Donald Miller

Abstract

We present measurements of the mean Lyman-α transmitted flux at $4.9 < z < 6.1$ along the lines of sight of 92 quasars at $z_{\rm source} > 5.7$, the largest sample assembled to date by a factor of three. This enables us to sample cosmic variance at those redshifts more robustly than ever before. The spectra used here were obtained by the SDSS, DES-VHS and SHELLQs collaborations, drawn from the ESI and X-Shooter archives, reused from previous similar studies or observed specifically for this study. We measure the effective optical depth of Lyman-α in bins of 10, 30, 50 and 70 cMpc $\rm h^{-1}$ and introduce a new way to treat upper limits on flux. We check the consistency of our results with the samples of Fan et al. (2006) and Becker et al. (2015b) via bootstrap re-sampling. We confirm the existence of a tail of high-opacity lines of sight at $5.1 < z < 5.3$, $5.3 < z < 5.5$ and $5.5 < z < 5.7$. We find slightly less opacity at $5.7 < z < 5.9$ compared to Becker et al. (2015b), and provide measurements for the redshift range $5.9 < z < 6.1$ for the first time. Comparing our results with predictions from cosmological simulations, we find strong evidence against a uniform UVB. While the models of Chardin et al. (2015) and Keating et al. (2017) both ease the tension without removing it, the former is a better fit to the observed spread of opacities. We briefly look at the statistics of Lyman-α transmission spikes in the most high-quality spectra and compare the results with models. Our sample of 92 $z > 5.7$ quasar spectra opens many avenues for future study of the reionisation epoch.

V.1 Introduction

The first billion years of the Universe are currently a frontier of late-time cosmology both observationally and theoretically. During this stretch of time the first stars and galaxies assembled from the primordial gas left behind by reheating, and the atomic hydrogen permeating the early Universe became ionised. It is known that this transition, called reionisation or Cosmic Dawn, was largely completed by redshift six. The precise timing and topology of reionisation are strongly influenced by the processes at work in the first galaxies and active galactic nuclei (AGN) as well as the large–scale structure of the early intergalactic medium (IGM).

Quasars located at $z > 6.0$ have proven useful tools in obtaining information about reionisation due to their high intrinsic luminosities and prominent Lyman-α emission lines. These properties have yielded results on multiple fronts, from measuring the sizes of quasar proximity zones across time, which are expected to diminish with increasing IGM neutrality and decreasing AGN lifetime (e.g. Fan et al., 2006; Carilli et al., 2010; Keating et al., 2015) to constraining enrichment processes by probing the cosmic abundances of excited metallic ions intervening along the lines of sight (e.g. Ryan-Weber et al., 2009; D'Odorico et al., 2013; Becker et al., 2015a; Chen et al., 2016; Bosman et al., 2017). The Lyman-α forest extending bluewards of 1215Å in the quasar rest frame is of particular interest to reionisation as it is able to resolve individual filaments and voids at $z < 5.0$. The absorption features from these structures start to merge above that redshift, and eventually complete absorption is reached once the IGM reaches average hydrogen neutral fractions of $f \geq 0.1$ per cent (Gunn & Peterson, 1965). The characterisation of Lyman-α opacity across redshift has become a powerful probe of models of reionisation, as the amount of residual transmission is sensitive to the nature of the UV sources, the thermal state of the IGM, and the large-scale clustering of sources among other factors (e.g. Wyithe & Bolton 2011; McQuinn et al. 2011; Davies et al. 2017).

Lyman-α transmission along QSO lines of sight is often quantified by an "effective optical depth": $\tau = -\ln(\langle F/F_0 \rangle)$, where $\langle F \rangle$ is the average flux over a stretch of the Lyman-α forest, and F_0 is the unabsorbed continuum. (Fan et al., 2006). The first studies of the optical depth distribution restricted themselves to small samples of high sensitivity data, obtained as part of a single high-z quasar survey or by the authors themselves (Becker et al., 2001; Fan et al., 2006; Becker et al., 2015b). Early measurements pointed to a large scatter in transmission along lines of sight potentially incompatible with fluctuations in the density field alone (Lidz et al., 2006; Bolton & Haehnelt, 2007b) but small sample sizes did not allow for in-depth study of the effect (Lidz et al., 2007; Mesinger, 2010). Becker et al. (2015b) discovered a ~ 110 cMpc Lyman-α trough extending down to $z = 5.4$ and demonstrated that its existence is incompatible with a smooth ultra-violet background (UVB). The discovery prompted a flurry of new reionisation models (see Chapter I.4.1). Their sample included 33 quasar spectra, a significant increase from previous studies, but still only a fraction of the more than 200 quasars now known at $z > 5.7$. In this paper, we gather 92 spectra of $z > 5.7$ quasars, nearly tripling upon the previously largest sample.

We present updated measurements of the Lyman-α opacity distribution function (PDF) for the redshift range $4.9 < z < 6.1$. The number of known quasars at $z > 5.9$ is increasing rapidly due to high redshifts searches by the Dark Energy Survey (DES; Reed et al., 2015), Subaru High-z Exploration of Low-Luminosity Quasars (SHELLQs; Matsuoka et al., 2016), PANSTARRS (Kaiser et al., 2010), VISTA Kilo-Degree Infrared Galaxy (VIKING; Venemans et al. 2013; Carnall et al. 2015) survey, the Canada-France High-redshift Quasar Survey

(CFHQS; Willott et al., 2007) and UKIDSS (Venemans et al. 2007; Mortlock et al. 2009, 2011) as well as the completion of the search for high redshift quasars in the Sloan Digital Sky Survey (SDSS, York et al., 2000, Jiang et al., 2016b). We provide smoother constraints on the Lyman-α PDF with a better handle on cosmic variance.In addition, we are able to push the measurement of Lyman-α opacity up to $z = 6.1$ for the first time.

The chapter is structured as follows. In Section 2 we present our sample of 92 quasars and present four previously unpublished spectra, briefly discussing the properties of our sample. Our methods for measuring Lyman-α opacity distributions are presented, and compared to previous studies in Section 3. Challenges in dealing with the wide range of spectral resolutions and signal to noise ratios (SNRs) across our sample are discussed. Section 4 gives our results spanning the redshift range $4.9 < z < 6.1$ using multiple ways of accounting for the inhomogeneous quality of the data and non-detections of transmitted flux. These results are confronted with predictions from numerical models and discussed in Section 5. Section 6 briefly discusses the use of Lyman-α transmission spike statistics as an extra tool for testing models of reionisation. The results are summarized in Section 7 and extra figures, including a mosaic of the entire sample, are shown in Section 8. Throughout the paper we use $(\Omega m, \Omega\Lambda, h) = (0.308, 0.692, 0.678)$ (Planck Collaboration et al., 2016a) and quote comoving distances in units of Mpc h^{-1}

V.2 Data

Our final sample consists of 92 quasar spectra observed at $z > 5.7$ over the last 11 years. Out of these objects, 9 are discovery spectra from the SHELLQs survey, 18 were discovered by DES-VHS (out of which 9 are currently unpublished), 25 are SDSS discovery spectra, 13 are new reductions of archival data, 23 were obtained in previous studies on Lyman-α transmission, and 4 are new to this work. Nine different optical spectrographs were used to obtain the data: ESI, X-Shooter, GMOS, MagE, EFOSC, FOCAS, MMT RCS, HIRES and LBT MODS. The following sections describe the make-up of the sample in more detail. Table V.2 details the provenance of each spectrum used in this work.

V.2.1 SDSS quasars

The SDSS is a sky survey over 14,555 deg^2 which provided imagining in the *ugriz* photometric bands as well as spectroscopic follow-up using a 2.5m dedicated telescope locate at Apache Point Observatory (Fukugita et al., 1996; Hogg et al., 2001). Here we briefly outline the detection procedure of quasars in the SDSS (see Jiang et al. 2016b for a more in-depth summary). Candidates are selected in the first step as drop-outs with no detections in the *ugr* photometric bands and with colors in excess of $i_{AB} - z_{AB} > 2.2$. After quality cuts, follow-up photometry is obtained in the near infrared (IR) and a second cut $z_{AB} - J < 0.5 + 0.5\Delta_{i-z}$ is imposed (e.g. Fan et al. 1999). Alternative colour cuts are used in deeper areas of the survey near the Galactic cap (Jiang et al., 2008, 2009) and in regions scanned two or more times (Jiang et al., 2015).

Confirmation spectra of the quasar candidates are typically obtained obtained with the Red Channel Spectrograph (RCS) on the 6.5m Multiple Mirror Telescope (MMT) or Double Spectrograph on the Hale 5.1m telescope (DBSP) (e.g. Jiang et al., 2016b), and in one occasion with the Multi-Object Double Spectrographs for the Large Binocular Telescope on Mt. Graham in southeastern Arizona (LBTMODS, Pogge et al. 2012). Additional near-IR spectra taken for

QSO	z_{em}	Instrument	SNR	Survey	Notes	Discovery ref.	Spectrum ref.
J1120+0641	7.0842	X-Shooter	35.0	UKIDSS		(1)	(26)
J0020-????	6.91	EFOSC	3.0	DES-VHS		(24)	–
J2348-3054	6.90	X-Shooter	13.4	VIKING	new reduction	(29)	(30)
J1205-0000	6.8	FOCAS	3.5	SHELLQs		(20)	–
J0420-????	6.75	GMOS	6.3	DES–VHS		(24)	–
J0224-????	6.75	GMOS	9.3	DES–VHS		(24)	–
J0224–4711	6.50	GMOS	6.5	DES–VHS		(2)	–
J0210–0456	6.44	ESI	5.3	CFHQS	new reduction	(3)	(32)
J2329–0301	6.43	ESI	6.5	CFHQS	new reduction	(11)	(25)
J1148+5251	6.419	HIRES	29.7	SDSS		(4)	(22)
J2236+0032	6.4	FOCAS	2.1	SHELLQs		(20)	–
J0859+0022	6.39	FOCAS	2.8	SHELLQs		(20)	–
J1152+0055	6.37	FOCAS	3.1	SHELLQs		(20)	–
J1148+0702	6.339	MMT	3.4	SDSS		(5)	–
J0100+2802	6.30	X-Shooter	85.2	SDSS	new spectrum	(6)	this paper
J1030+0524	6.28	X-Shooter	28.0	SDSS	new reduction	(7)	(22)
J0050+3445	6.25	ESI	24.4	CFHQS		(3)	(23)
J0143–5545	6.25	GMOS	3.5	DES–VHS		(2)	–
J0323–4701	6.25	EFOSC	12.5	DES–VHS		(2)	–
J0330–4025	6.25	EFOSC	12.1	DES–VHS		(2)	–
J1623+3112	6.247	ESI	16.1	SDSS		(8)	(22)
J2325-????	6.23	EFOSC	1.8	DES–VHS		(24)	–
J0410–4414	6.21	EFOSC	12.7	DES–VHS		(2)	–
J0227–0605	6.20	ESI	7.5	CHFQS	new reduction	(27)	(25)
J1048+4637	6.198	HIRES	29.2	SDSS		(4)	(28)
J2232+0012	6.18	FOCAS	2.1	SHELLQs		(20)	–
J1609+3041	6.16	MMT	6.1	SDSS		(5)	–
J2229+1457	6.15	ESI	6.0	CHFQS	new reduction	(3)	(25)
J1250+3130	6.13	ESI	26.2	SDSS		(9)	–
J0033–0125	6.13	ESI	6.1	CHFQS	new reduction	(11)	(25)
J1319+0950	6.132	X-Shooter	96.8	UKDISS/SDSS		(10)	(23)
J1509–1749	6.12	X-Shooter	88.9	CFHQS		(11)	(22)
J2315–0023	6.117	ESI	29.8	SDSS		(12)	(23)
J0454–4448	6.10	MagE	5.8	DES		(21)	–
J0109–5423	6.10	EFOSC	4.2	DES–VHS		(24)	–
J2216-0016	6.10	FOCAS	2.4	SHELLQs		(20)	–
J1602+4228	6.09	MMT	33.3	SDSS	new reduction	(8)	(25)
J0303–0019	6.078	ESI	8.0	SDSS	new reduction	(12)	(32)
J0353+0104	6.072	ESI	80.7	SDSS		(12)	(23)
J0420–4453	6.07	EFOSC	6.4	DES–VHS		(2)	–
J0842+1218	6.069	ESI	21.4	SDSS		(13)	(22)
J2054–0005	6.062	ESI	39.5	SDSS		(12)	(23)
J1630+4012	6.058	MMT	17.0	SDSS		(4)	(22)
J1207+0630	6.040	MMT	7.3	SDSS		(13)	–
J1641+3755	6.04	ESI	9.0	CHFQS	new reduction	(11)	(25)
J0408–5632	6.03	EFOSC	4.3	DES–VHS		(2)	–
J0159–3633	6.02	EFOSC	2.7	ATLAS/DES–VHS		(14)	this paper
J1257+6349	6.02	MMT	6.1	SDSS		(13)	–
J1306+0356	6.016	X-Shooter	55.8	SDSS	new reduction	(7)	(22)
J1137+3549	6.01	ESI	31.7	SDSS	new spectrum	(9)	this paper
J2228+0128	6.01	FOCAS	2.2	SHELLQs		(20)	–
J1207-0005	6.01	FOCAS	3.1	SHELLQs		(20)	–
J2310+1855	6.003	LBTMODS	17.9	SDSS		(5)	–
J0818+1722	6.0	HIRES	39.8	SDSS		(9)	(28)
J2250–5015	6.00	EFOSC	5.1	DES–VHS		(2)	–
J2356+0023	6.00	MMT	3.9	SDSS		(16)	–
J0131–3841	6.00	EFOSC	3.9	DES–VHS		(24)	–

Table V.1 – Current list of quasars. References given in caption of Fig V.2.

QSO	$z_{\rm em}$	Instrument	SNR	Survey	Notes	Discovery ref.	Spectrum ref.
J0028+0457	5.99	MMT	6.0	SDSS		(13)	–
J0532–4401	5.97	EFOSC	2.3	DES–VHS		(24)	–
J0841+2905	5.96	ESI	11.2	SDSS		(15)	(22)
J0433–5153	5.96	EFOSC	2.1	DES–VHS		(24)	–
J0122-????	5.96	EFOSC	3.0	DES–VHS		(24)	–
J1202–0057	5.93	FOCAS	2.2	SHELLQs		(20)	–
J0008–0626	5.929	MMT	4.4	SDSS		(13)	–
J1411+1217	5.927	ESI	15.9	SDSS		(8)	(22)
J0148+0600	5.923	X-Shooter	128.0	SDSS		(13)	(23)
J2053+0047	5.92	MMT	5.2	SDSS		(16)	–
J1335+3533	5.901	ESI	16.3	SDSS		(9)	(22)
J0835+3217	5.89	MMT	2.0	SDSS		(5)	–
J2119–0040	5.87	MMT	4.0	SDSS		(5)	–
J2307+0031	5.87	MMT	3.6	SDSS		(5)	–
J0850+3246	5.867	MMT	3.7	SDSS		(13)	–
J0203+0012	5.86	ESI	17.4	UKIDSS/SDSS		(12)	(23)
J1403+0902	5.86	MMT	7.3	SDSS		(13)	–
J0005–0006	5.850	ESI	28.8	SDSS	new reduction	(8)	(32)
J1243+2529	5.85	MMT	4.1	SDSS		(5)	–
J0840+5624	5.844	ESI	17.6	SDSS		(9)	(22)
J1215+0023	5.84	X-Shooter	19.6	VIKING+KiDS	new reduction	(31)	(30)
J1436+5007	5.83	MMT	3.2	SDSS		(9)	–
J0239–0045	5.82	MMT	4.9	SDSS		(16)	–
J0836+0054	5.810	X-Shooter	93.4	SDSS	new reduction	(7)	(22)
J1143+3808	5.81	MMT	10.5	SDSS		(5)	–
J2147+0107	5.81	MMT	4.1	SDSS		(16)	–
J0002+2550	5.8	HIRES	71.7	SDSS		(8)	(28)
J0810+5105	5.80	MMT	10.0	SDSS		(5)	–
J1044–0125	5.80	MagE	35.7	SDSS		(17)	(22)
J1545+6028	5.78	LJT G5	?	SDSS–WISE	no noise array	(18)	–
J0129–0035	5.779	MMT	3.7	SDSS		(16)	–
J0927+2001	5.772	X-Shooter	73.7	SDSS	new spectrum	(9)	this paper
J2325+2628	5.770	MMT	3.8	SDSS–WISE		(18)	–
J1420–1602	5.73	MagE	29.4	PANSTARRS		(19)	(22)
J1621+5155	5.71	MMT	10.3	SDSS		(5)	–

Table V.2 – Current list of quasars, continued. Quasar names including question marks are not public yet. References: (1) Mortlock et al., 2011; (2) Reed et al., 2017; (3) Willott et al., 2010; (4) Fan et al., 2003; (5) Jiang et al., 2016b; (6) Wu et al., 2015; (7) Fan et al., 2001; (8) Fan et al., 2004; (9) Fan et al., 2006; (10) Mortlock et al., 2009; (11) Willott et al., 2007; (12) Jiang et al., 2008; (13) Jiang et al., 2015; (14) Carnall et al., 2015; (15) Goto, 2006; (16) Jiang et al., 2009; (17) Fan et al., 2000; (18) Wang et al., 2016; (19) Morganson et al., 2012; (20) Matsuoka et al., 2016; (21) Reed et al., 2015; (22) McGreer et al., 2015; (23) Becker et al., 2015b; (24) Reed in prep.; (25) KOA 2017; (26) Bosman et al. 2017; (27) Willott et al. 2009; (28) Becker et al. 2006; (29) Venemans et al. 2013; (30) X-Shooter archives; (31) Venemans et al. 2015; (32) Eilers et al. 2017

some objects do not cover the range 7500Å – 10,000Å required for coverage of Lyman-α at $5.3 < z < 7.0$ and are not used in this work (e.g. Jiang et al., 2007, Simcoe et al., 2011).

Jiang et al. (2016b) presented the 52 final quasars discovered by the SDSS, all of which are included in this work. Out of those, 29 have been re-observed with better data quality since their discovery, while 23 have not. The discovery spectra for those 23 are included in our sample. 10 of those objects were reported in Jiang et al. (2016b), 6 objects in Jiang et al. (2015), 5 objects in Jiang et al. (2009) and 2 objects in Fan et al. (2006). Two additional quasars discovered by combining SDSS photometry with WISE data are also included (Wang et al., 2016).

V.2.2 DES and DES–VHS quasars

The Dark Energy Survey (DES) covered an area of 5000 deg^2 in the southern hemisphere in visible imaging. It employed the dedicated Dark Energy Camera (DECam) on the Blanco 4m telescope, Cerro Tololo (The Dark Energy Survey Collaboration, 2005). The first high-z quasar discovered in DES was presented in Reed et al. (2015). Quasar candidates were selected using a drop-out technique similar to the SDSS procedure described above, this time with the condition $i_{DES} - z_{DES} > 1.694$. In addition the DES survey includes the *Y* red optical band, allowing a more efficient removal of red dwarves from the sample via a constraint on the quasar continuum of $z_{DES} - Y_{DES} < 0.5$. In Reed et al. (2017), eight additional quasars were detected by combining DES data with infrared observations of overlapping footprint of the VISTA Hemisphere Survey (VHS; McMahon et al. 2013). Nine additional objects have been discovered in the same way since (Reed et al, in prep).

Spectroscopic confirmation of the candidates was conducted either with the ESO Faint Object Spectrograph and Camera (EFOSC, Buzzoni et al. 1984) or the Gemini Multi-Object Spectrographs (GMOS, Hook et al. 2004), with some objects subsequently observed in higher quality with the Magellan Echellette (MagE, Marshall et al. 2008). The best quality spectrum for each of the 18 DES–VHS quasars was chosen as shown in Table V.2.

V.2.3 SHELLQs quasars

The Subaru High-z Exploration of Low-luminosity Quasars (SHELLQs, Matsuoka et al. 2016) is a new imaging survey utilising the Hyper Suprime-Cam on the Subaru 8.2m telescope (Miyazaki et al., 2012). A search for quasars has currently been conducted over an area of 430 deg^2. The SHELLQs project aims to obtain deeper exposures in the *grizy* bands compared to SDSS and DES, leading to the discovery of 33 faint $z > 5.7$ quasars so far (Matsuoka et al., 2016, 2017). In this work, we include the first nine SHELLQs quasar spectra presented in Matsuoka et al. (2016). The confirmation spectra for these objects were taken with the Faint Object Camera and Spectrograph on the Subaru telescope (FOCAS, Kashikawa et al. 2002) as described in the discovery paper.

V.2.4 Other quasar spectra

In this work, we re-use 22 quasar spectra presented in previous investigations of Lyman-α opacity. McGreer et al. (2011) and McGreer et al. (2015) conducted observations of 22 previously known quasars with the MagE, MMT and the X-Shooter instrument on Cassegrain UT2 (Vernet et al., 2011).

Object	z	Instrument	Date	Exposure Time (s)	Slit Width (arcsec)	Seeing (arcsec)
SDSS J0100+2802	6.3	X-Shooter	23rd Oct 2015	1800	1.00	0.80
ATLAS J0159-3633	6.02	EFOSC	18th Oct 2015	1800	1.50	0.72
SDSS J1137+3549	6.01	ESI	18th Mar 2016	3000	1.00	0.80
SDSS J0927+2001	5.772	X-Shooter	13th Jan 2010	1800	1.00	0.77

Table V.3 – New observations of four $z > 5.7$ quasars which are were not presented in previous work.

We are making use of 12 of those observations as indicated in Table V.1. Similarly, Becker et al. (2015b) provided 7 spectra taken on the Echellette Spectrograph and Imager (ESI) on the Keck II telescope (Sheinis et al., 2002), and X-Shooter, of quasars at $z = 6.25, 6.13, 6.12, 6.06, 6.05$ and $z = 5.98$. Three additional spectra were first presented in Becker et al. (2006). The quasars followed up in those papers were initially discovered by various surveys including the UKIRT Infrared Deep Sky Survey (UKIDSS, Lawrence et al. 2007), the Canada-France High-z Quasar Survey (CFHQS, Willott et al. 2007), SDSS, and the Panoramic Survey Telescope and Rapid Response System (Pan-STARRS, Morganson et al. 2012). In addition, we also include a 30h X-Shooter spectrum of ULAS J1120+0641 at $z = 7.08$, first presented in Bosman et al. (2017).

V.2.5 New reductions

After a proprietary period of 18 months, raw data obtained with ESI is made publicly available through the Keck Observatory Archive (KOA[1]). In this work, we use 6 ESI spectra of quasars at $5.85 < z < 6.43$ re-reduced from raw data obtained from the KOA. Our ESI pipeline is the same as described in Subsection V.2.6.

Similarly, raw data from the X-Shooter instrument is released online 12 months after observations and can be obtained through the X-SHOOTER Raw Data Query Form[2]. We present reductions of two quasars observed by X-Shooter: J1215+0023, a $z = 5.84$ object discovered by the VIKING+KiDS collaboration (Venemans et al., 2015), was observed in 2014 and 2015 as part of program 093.A-0851(A) (PI: Verdoes Kleijn). Quasar J2348-3054, at $z = 6.90$, was discovered in the VIKING survey Venemans et al. (2013) and observed in 2011 and 2012 in program 087.A-0890(A) (PI: De Rosa). We reduced both of those archival spectra using a custom pipeline described below.

Finally, we re-reduced the spectra of five quasars which have been previously published. X-Shooter spectra of the quasars J0836+0054 and J1030+0524 were introduced in McGreer et al. (2015), but here we use our own reduction of the raw X-Shooter files instead, in an attempt to improve data quality. ESI spectra of quasars J0210-0459, J0303-0019, and J0005-0006 were part of an observing run (find ID) in October 2010 (PI: Sargent) and have been previously used in e.g. Eilers et al. (2017). Here we use our own reductions of the raw ESI data.

[1] `https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOAlogin`
[2] `http://archive.eso.org/wdb/wdb/eso/xshooter/form`

V.2.6 New spectra

We present 4 new observations of quasars which were carried out on the ESI, X-Shooter and EFOSC instruments as detailed in Table V.2.

An X-Shooter spectrum was obtained in January 2010 of quasar SDSS J0927+2001 at $z = 5.772$ but had not previously been utilised. The reduction procedure is the same as the one presented in e.g. Becker et al. (2015b). The spectrum was extracted optimally (Horne, 1986) using 10 km s^{-1} bins after being flat-fielded and sky-subtracted following Kelson (2003). Custom telluric absorption routines were used as presented in Becker et al. (2012).

A 3000s ESI spectrum of the $z = 6.01$ quasar SDSS J1137+3549 was obtained by the author on the 18th of March, 2016.

ALTAS J0159-39633, at $z = 6.02$, was accidentally re-discovered during spectroscopic follow-up of DES quasar candidates. It was observed with the EFOSC instrument at the same time as the quasars presented in Reed et al. (2017).

A deep X-Shooter spectrum of J0100+2802 was obtained by George Becker and Max Pettini over 2015 and 2016 as part of the 13 hours program 096.A-0095(A). Here we make use of only one exposure of 1800s of the object, which is nevertheless a great improvement upon the previous LBT-MODS spectrum of the quasar.

V.2.7 Sample properties and notes on individual objects

Our 92 quasars have source redshifts in the range $5.70 < z < 7.08$ with a peak at $z \sim 6.0$ and a distribution as shown in Fig V.2. The usable portion of the spectra is $1041 < \lambda_{\rm rest} <\sim 1178$Å, resulting in a redshift coverage shown in Fig V.3 with up to 82 lines of sight covering the region $5.4 < z < 5.5$. These distributions vary slightly depending on the choice of proximity-zone cut-off.

Redshifts for the objects in our sample are based on best available estimates using nebular lines (such as [CII]) whenever possible. See Jiang et al. (2016b) for the origin of the measured redshifts of SDSS quasars. Since we are primarily concerned with measuring Lyman-α opacity in the IGM, accurate values for the AGN redshifts are only relevant to the exclusion of AGN proximity zones from the analysis. In this work we are not attempting to measure the evolution of quasar proximity zone length across redshift. We use a fix cut–off for the end of the proximity zone of $\lambda_{\rm end,\ prox} = 1178$ Å after checking that this is a reasonable choice and that more stringent criteria do not affect the results. This analysis is presented in Section V.3.1.

We measure the signal to noise ratio (SNR) for the quasars in the sample listed in column four of Table V.2. The measurement is complicated by the disparate resolutions and redshifts of the quasar spectra, since common sky lines fall at different rest wavelengths across the sample. To measure SNR, we first normalise the spectrum by a power law as described above, then keep the pixels located at $1275 < \lambda < 1285$ Å which are not affected by sky lines. This wavelength range covers a portion of quasar spectra minimally affected by broad emission lines. Measuring SNR over this range provides an estimate of the difficulty of resolving, and therefore fitting, the quasar continuum. We present here the mean SNR per 60 km s^{-1}. The mean flux is computed over the 10 Å range measured above, then rescaled according to the spectral pixel size following,

$$\text{SNR} = \left\langle \frac{F}{\epsilon} \right\rangle \cdot \sqrt{N_{60}} \tag{V.1}$$

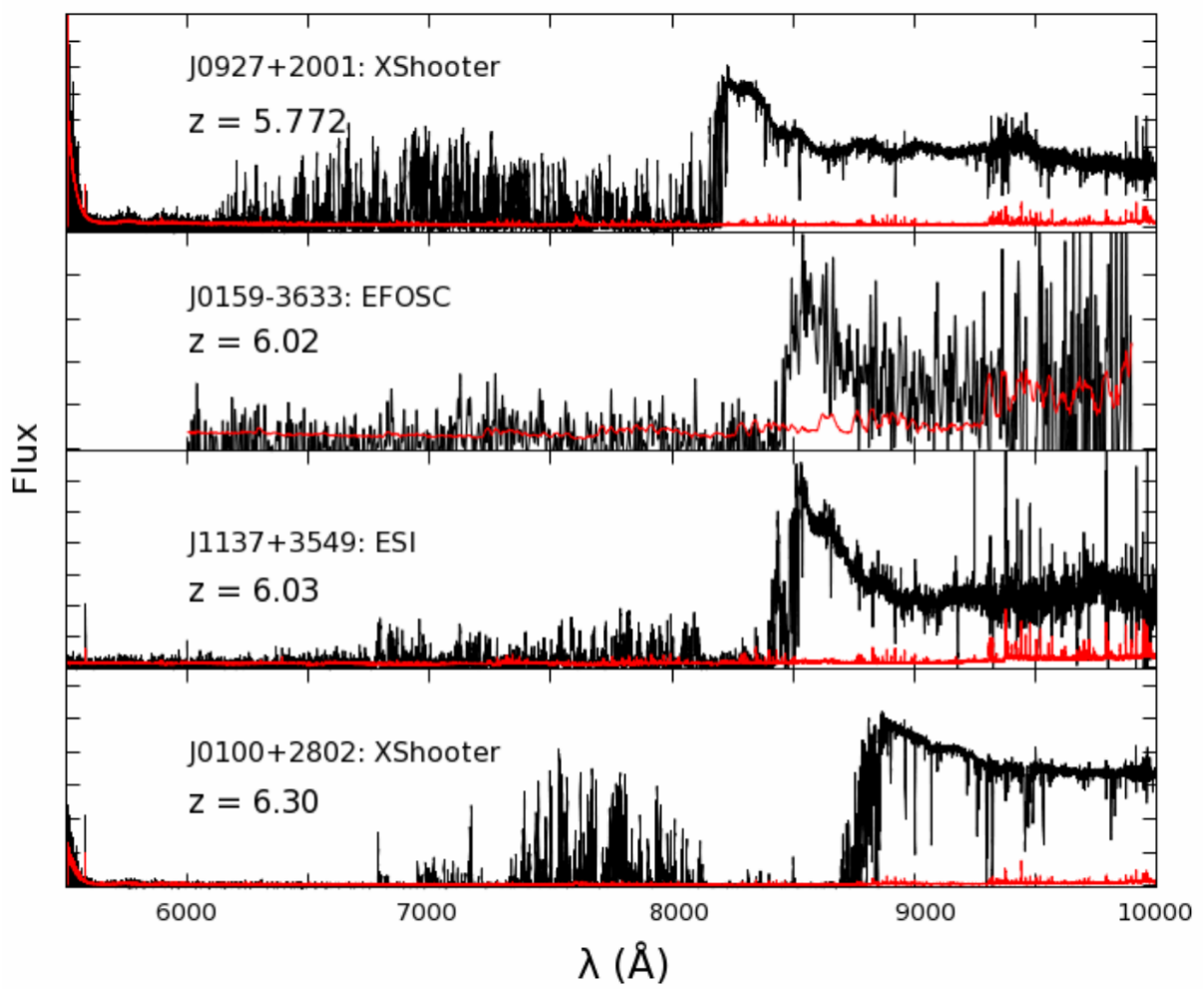

Figure V.1 – New quasar spectra used in this work.

where ϵ is the error and N_{60} the number of pixels per 60 km s^{-1} interval. This is computed using $N_{60} = 60.0/\Delta v$, where Δv is the pixel size in km s^{-1}, related to the resolution via the speed of light c: $\Delta v/c = \Delta\lambda/\lambda$. For spectrographs providing a fixed $\Delta\lambda$ rather than a constant resolution, Δv is measured at $\lambda = 8000$ Å. While this is not the only way of homogeneously measuring spectral SNR, it is sufficient for our purposes to discriminate between data quality regardless of resolution, and has the advantage of being invariant under re-binning of the spectra. In particular, we note that rescaling the SNR by pixel size, instead of resolution element, might introduce correlations and lead to a slight underestimate of the actual noise. However, the information relating to resolution in the raw data was not available or reliable for all spectra. The values obtained range from SNR$= 1.8$ for J2325-5229, a DES–VHS quasar with continuum emission barely above the detection threshold, to SNR$= 96.8$ in a deep X-Shooter exposure of J1319+0959 first used in Becker et al. (2015b).

Out of the quasars in our sample, two show the characteristic features of broad absorption line spectra (BAL) and one shows clear contamination by a DLA. Below we discuss these three objects in turn. We note that the fraction of quasar spectra displaying these features is lower than measured at lower redshift. This could be due to the selection techniques employed to discover high redshift quasars, as the presence of a BAL feature diminishes the photometric

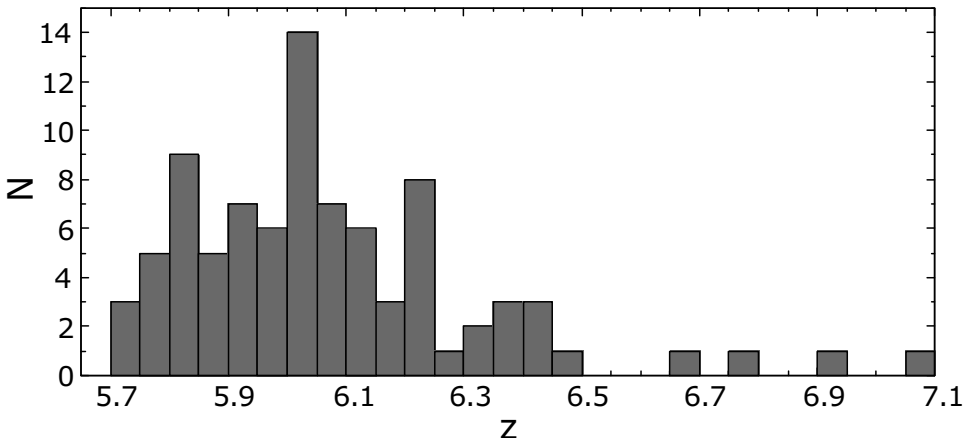

Figure V.2 – Redshift distribution of the quasars included in our sample.

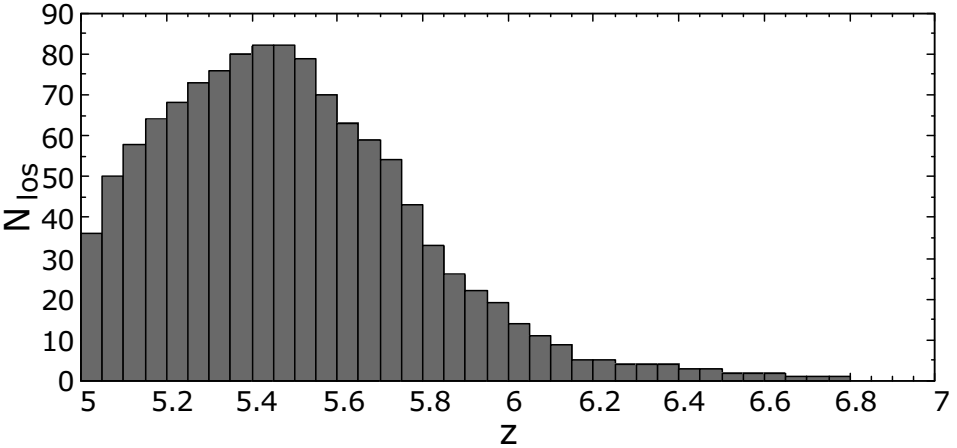

Figure V.3 – distribution of the total number of lines of sight covering a particular redshift.

colours most commonly used to select quasar candidates as drop-outs. It is also likely that some contamination by DLAs has gone unresolved in our sample. Due to the saturation of the Lyman-α forest, the only way to detect and remove DLAs from the sample is by detecting associated metal absorption at the DLA redshift. The wavelength coverage means that low-ionisation line complexes such as Mg II are unfortunately not visible for the redshifts of interest here, and the quality of the spectra is insufficient to detect typical C IV absorption systems in the majority of cases. Only deep X-Shooter spectra provide sufficient coverage and sufficient SNR to completely remove DLA contamination at $z > 5.5$.

J0859+002, at $z = 6.39$, displays very strong Lyman-α and broad N V emission identified by Matsuoka et al. (2016) but unfortunately shows a saturated Lyman-α absorption line at +3640 km s^{-1} from the peak of Lyman-α . We cannot exclude that N V emission is instead abnormally strong compared to the continuum, which is barely detected beyond $\lambda = 1250$Å. This object also displays a well-defined proximity zone. There is no detection of any transmitted Lyman-α flux. We decide to exclude this object from our analysis.

Quasar J2325-5229, at $z = 6.23$, is similarly an unpublished DES quasar showing a strong Lyman-α absorption feature +3320 km s^{-1} from the peak of Lyman-α emission. No Lyman-α transmission is detected, although the large errors and the short available continuum make a detection unlikely in this spectrum. We also exclude this object from our analysis.

Finally, we noticed that J2053+0047 at $z = 5.92$ seems to posses a particularly opaque Lyman-α trough over the ranges $4.9 < z < 5.1$ and $5.1 < z < 5.3$, with $\tau_{\rm eff} = 3.66$ and 6.34 respectively. This would make it the most opaque line of sight at those redshifts by some distance, and carry significant implications for models of reionisation. Upon closer inspection, we however find a blended, saturated C IV doublet intervening along the line of sight at $z = 5.06$. This leads us to suspect the presence of a DLA along the line of sight and to exclude this object from the analysis. Numerical simulations do not resolve the effects of self-shielding at work in DLAs.

V.3 Methods

The spectra are first normalised by fitting a power law to the unabsorbed continuum. The portion being fitted extends from 1270 Å–1450 Å in the rest frame of the quasar; the range 1270 Å– 1350 Å is used instead when the spectral coverage stops short of 1500 Å to avoid portions of the spectrum affected by the falling response of the instrument. This is the case for instance for spectra of $z > 5.7$ quasars taken with the MMT or MagE, whose coverage extends to 10,000Å but the response of which decays significantly from $\gtrsim 9700$Å. Pixels affected by sky lines are excluded and a first power law (PL) fit is made, from which we then exclude any pixels for which $|F_{\rm QSO} - F_{\rm PL}| > 2\epsilon_{\rm QSO}$ – a process referred to as sigma clipping. The remaining flux is then fit with a power law again, and the process repeated a second time with a deviation coefficient of 1.5 to ensure convergence. Finally, the full flux array is divided pixel-by-pixel by the best fit power law function thus obtained.

Three objects which displayed too little continuum (due to a combination of high redshift and spectrograph used) are excluded from the analysis since no satisfactory estimate of the continuum could be obtained. The continua for all other objects were checked visually. We checked that the best fit power law parameters were robust to small changes in the fitting window bounds. The effects of window choices were run all the way through the analysis; we find

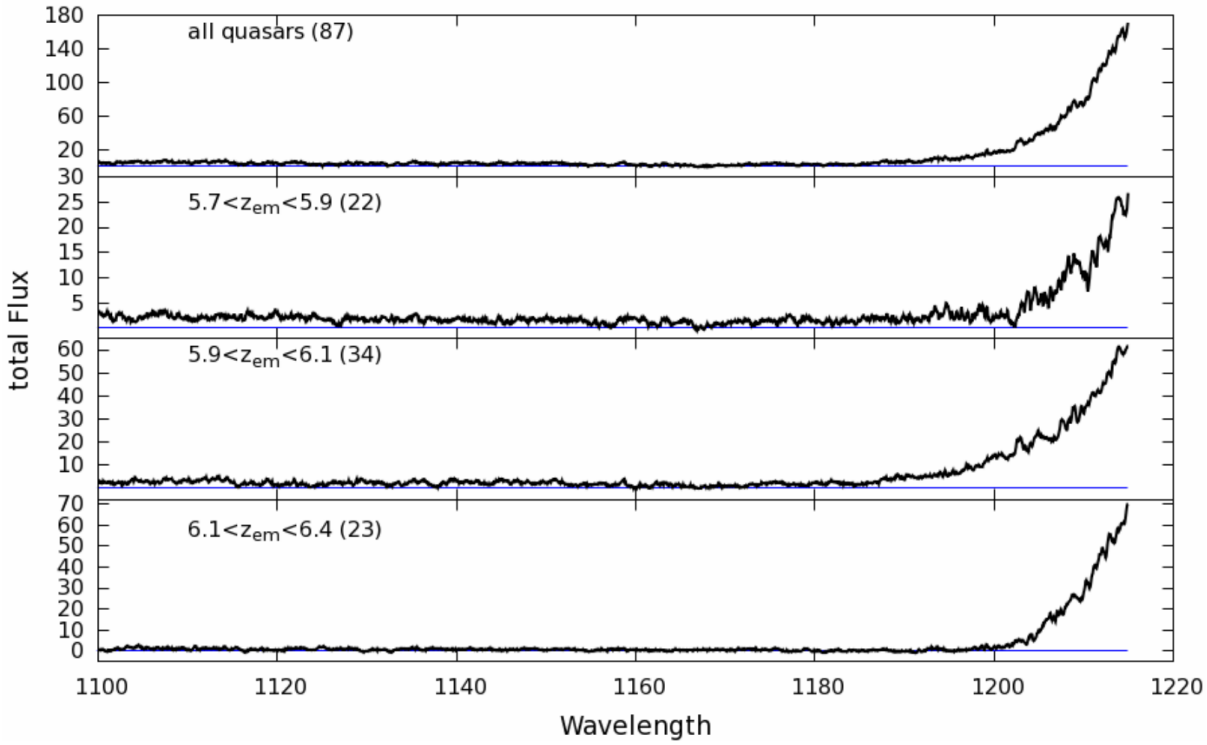

Figure V.4 – Stack of flux over 1100 Å–1215 Å of 87 $z > 5.5$ quasars. Spectra were normalised and sky lines masked before stacking. While the proximity zone transmission has fallen considerably by 1180Å, the exact end of the host quasar's influence on the flux is unclear. Five spectra are excluded as described in Section V.2.7. Panels show the evolution of the stacked spectrum over $5.7 < z < 5.9$, $5.9 < z < 6.1$ and $6.1 < z < 6.4$. The thin blue lines show a total flux of zero.

end effect on values of tau of magnitude $\Delta\tau \lesssim 10\%$. This effect can be seen in Fig.V.6, where the only differences between our measurements and those of Becker et al. (2015b) are due to small differences in the choices of continuum fitting. These errors are in all cases much smaller than the effect of cosmic variance.

We measure the average transmitted flux in windows of 50 comoving cMpc h^{-1} extending from the end of the quasar's proximity zone down to the onset of Lyman-β absorption (1041 Å in the rest frame). The average flux is transformed into opacity following $\tau = -\log(F)$ and associated to the redshift corresponding to the middle of that 50 cMpc h^{-1} region. The analysis is repeated for window sizes of 10, 30 and 70 cMpc h^{-1}.

We treat non-detections of transmitted flux in two different ways. First, following previous work, we take the upper limit on the flux to correspond to twice the error in the flux over the window. If individual peaks of transmission are detected at more than 2σ significance over that range, then we take the lower limit on the flux to be twice the flux in those peaks alone, following $F > F_{\text{peaks}} - 2\sigma_{\text{peaks}}$ where σ_{peaks} is the error over the wavelength covered by peaks (Becker et al., 2015b). This allows us to compare our results to the previous samples of Fan et al. (2006) and Becker et al. (2015b) which are a sub-set of our catalogue. The Lyman-α opacities in these two samples were not measured in identical ways, as the lengths of the excluded proximity zones and the details of the continuum fitting were subtly different. This might have resulted in a mild tension between the two samples, which we can now harmonize

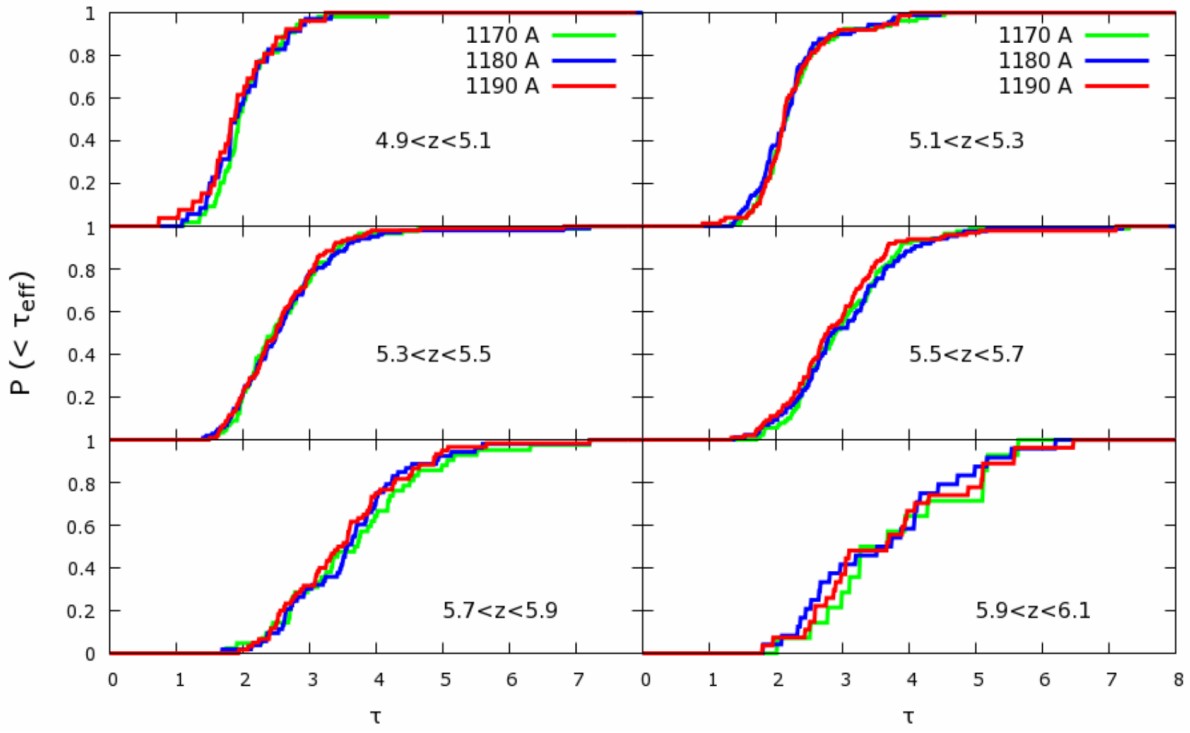

Figure V.5 – Effect of incrementally increasing the excluded proximity zone size on the Lyman-α CDF at various redshifts (see also Fig V.5). By ending the proximity zone at $\lambda = 1190$ Å, the average opacity is affected at $z < 5.3$ and $5.5 < z < 5.7$. However, no statistically significant difference is seen between a cut-off at $\lambda = 1180$ Å and the very conservative case $\lambda = 1170$ Å, at any redshift. Throughout this paper we adopt the traditional cut-off $\lambda = 1178$

by

1. doing a bootstrap re-sampling of our catalog to mimic previously used sample sizes, and

2. treating both samples identically through the same continuum fitting routine and same proximity zone exclusion technique.

Secondly, we treat upper limits in a more transparent way by plotting the most optimistic and pessimistic bounds on the cumulative distribution function (CDF). The optimistic bound is given, as above, by taking the average flux to be equal to two times the average error over the measurement window, i.e. just below detection threshold. However, different spectrographs with different exposure times will be sensitive to different thresholds: for instance, none of the data in our sample could measure an effective optical depth $\tau \geq 8.0$. To reflect this ambiguity, we attribute maximal opaqueness ($\tau \rightarrow \infty$) to all non-detections in order to obtain a 'maximally pessimistic bound' (so called due to the increased difficulty of reconciling this outcome with current reionisation models). The 'true' CDF will necessarily lie in-between these two extremes.

V.3.1 Proximity zone exclusion

We aim to measure the Lyman-α opacity of the IGM. In order to achieve this, we need to avoid any bias introduced by the proximity zones of the quasar sources: the region immediately surrounding the quasar where the ionisation of the gas is enhanced due to ionising radiation from the object itself. In the past, cut-offs for quasar proximity zones were chosen either on a case-by-case basis as the point where the quasar's Lyman-α flux had fallen below 10% of its peak value (Fan et al., 2006; Eilers et al., 2017) or chosen as a fixed value of 1176 Å determined based on a small sample of objects (Becker et al., 2015b). The former definition is not useful in the context of a sample containing spectroscopic data of varied resolution since clumps of neutral gas within the proximity zone might be resolved by some instruments but not others. A fixed-value cut-off is therefore more suited to a large dataset and facilitates future refinements of the measurements.

To ensure that the traditional value of 1176Å is sufficiently stringent as to remove all proximity zone influence, we first plot a stacked spectrum of our entire catalog in the range $1100 < \lambda < 1215$ Å (Fig V.5). The stacking was carried out by interpolating the spectra onto a common wavelength array after normalising them by a fitted power law and removing bad pixels. For the purposes of the stacking, objects with different measurements errors are not weighed accordingly. Some interesting features are visible, such as the slight increase in average opacity along lines of sight with redshift and the average power-law shape of the proximity zone. Based on the stack, we see that the proximity zone influences average transmitted flux at $\lambda > 1190$Å, but effects further away are unclear. To confirm and refine this, we incrementally increase the amount of excluded flux blueward of the Lyman-α line and compare the resulting CDFs. This is equivalent to restricting the measurement to quasars which are increasingly distant from the redshift of interest. The results are shown in Fig V.5 for several redshift bins are shown. The effect of the proximity zone flux is visible when a cut-off of $\lambda = 1190$ Å is used as the average opacity decreases. No statistically significant difference is seen between the $\lambda = 1180$ Å and the $\lambda = 1170$ Å case, indicating that the shape of the CDF has converged. We therefore stick to a value of the proximity zone end at $\lambda = 1178$ Å in the rest of the paper.

It is relevant to note that individual quasars showing anomalously long proximity zones (whether due to extreme bolometric luminosities, location in an under–dense IGM region, or chance) could be present in the sample and would not appear in the stacks in Fig V.5. The resulting contamination could potentially bias the average opacities to be too low. However, we do not find any such objects in the sample by visual inspection. Uncovering trends in quasar proximity zones is beyond the scope of this work, and it is enough for our purposes to confirm that no boost to average flux is seen at $\lambda < 1178$ Å in an averaged sense.

V.4 Results

V.4.1 Comparison to previous studies

As a first test of our procedure, we reproduce the cumulative PDF presented in Becker et al. (2015b) using the spectra of the 26 $z > 5.7$ quasars used in that work, which are a subset of our catalog. Figure V.6 presents the results at $5.5 < z < 5.7$ and $5.7 < z < 5.9$. The latter measurement makes exclusive use of those 26 quasars, and therefore any deviations are due to subtle differences in the continuum fitting – such as the precise wavelength ranges used and

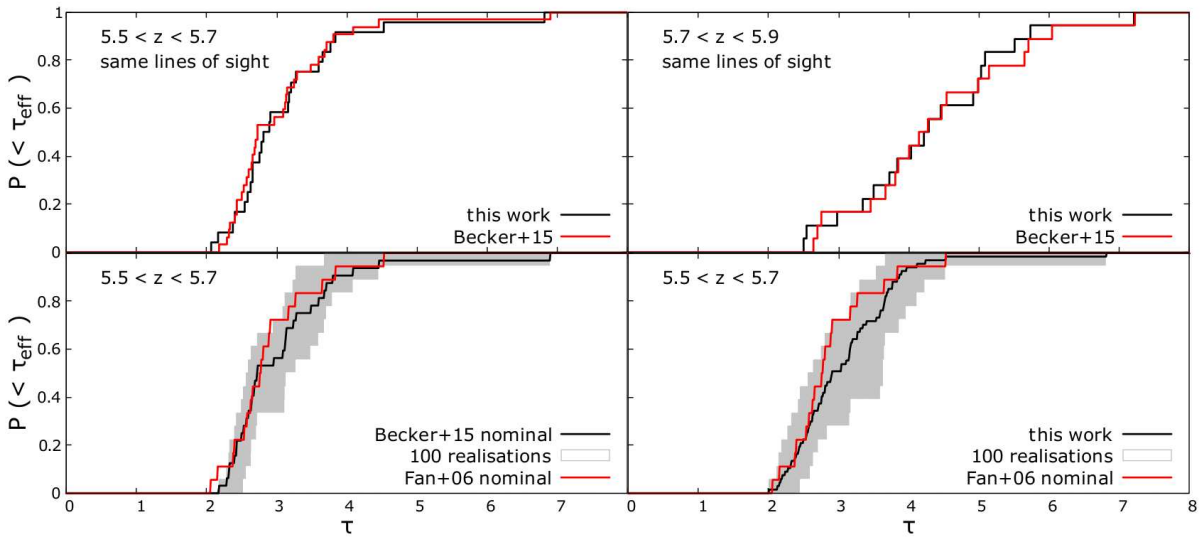

Figure V.6 – *Top:* Reproduction of the Becker et al. (2015b) opacity PDFs at $5.5 < z < 5.7$ (*left*) and $5.7 < z < 5.9$ (*right*). The same lines of sight are used, but 4 quasars at $z < 5.7$ are excluded from our work for the lower redshift case. Differences are accounted for by slight differences in the continuum fitting between the authors and previous measurements taken from Fan et al. (2006). *Bottom:* Reconciling the results of previous works by drawing a random sample of the same size used by Fan et al. (2006) from the sample of Becker et al. (2015b) (*left*) and this work (*right*) one hundred times. The shaded regions spans the obtained CDFs. The differences between previous results are consistent with random variance. All measurements are in bins of 50 cMpc h^{-1}.

the number of sigma-clipping iterations. In addition, the opacities of the spectra used in Fan et al. (2006) were not recomputed in Becker et al. (2015b), giving rise to a separate set of slight discrepancies. All our measurements of τ in 50 cMpc h^{-1} windows agree within error with those quoted in the original papers. The results at $5.5 < z < 5.7$ make use of 4 quasars at $z < 5.7$ which are not used in our sample. Nevertheless the shapes of the cumulative PDFs are in remarkable agreement, and the measurements of τ in individual windows all agree within error.

Next, we reconcile the results of Becker et al. (2015b) and Fan et al. (2006) by computing the CDF obtained from a random sub-sample of 19 quasars out of 33. One hundred such realisations are plotted in the bottom left panel of Figure V.6. We repeat the exercise using a random sampling of 19 out of our entire sample of 92 quasars while preserving the number of measurements of τ in the original work. In both cases we find the results are in statistical agreement, with the observed differences between the samples explained by random variation.

V.4.2 Lyman-α opacity CDF

Figure V.7 presents our updated results compared to the previous CDFs of Becker et al. (2015b) measured in the same away as those authors. Figure V.8 presents the same results, plotted to show the 'pessimistic' and 'optimistic' bounds described earlier. The results over $4.9 < z < 5.7$ are completely consistent with previous studies. We find a clear, well-defined tail of high-opacity ($\tau_{\rm eff} > 3$) lines of sight at redshift $z \sim 5.2$. This trend was already visible in the CDF

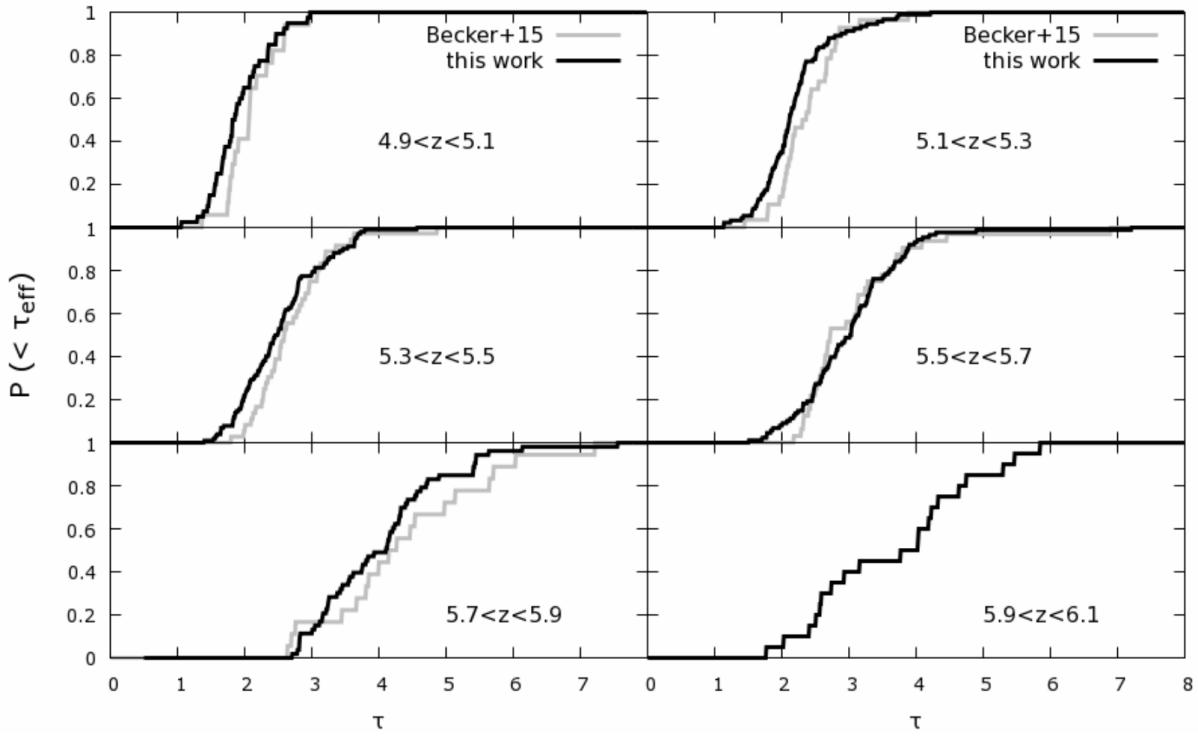

Figure V.7 – New results, obtained using the method described in Becker et al. (2015b). Only spectra of moderate or good quality (SNR>3.0) are used, and non-detections of transmitted flux are treated as a data point with value of twice the average error (see text). No significant discrepancy with previous results is found.

reported in Becker et al. (2015b).Roughly $\sim 20\%$ of lines of sight at $z = 5.2$ have opacities $\tau > 2.5$, which might pose problems for models of reionisation in which percolation of UV photons is expected to have happened long before this time.

At $z \sim 5.6$, we find a small but significant tail of transparent lines of sight, with roughly $\sim 20\%$ of measurements showing $\tau < 2.3$. This tail was not visible in the Becker et al. (2015b) sample as most of the relevant objects were not included. The two samples are consistent according to a KS test ($p > 0.10$). At $z \sim 5.8$, we find that opacities are slightly smaller than the ones previously reported. Again, the two samples are consistent according to a KS test ($p > 0.10$, Kolmogorov 1933). Our sample for $z \sim 5.8$ is of comparable size to Becker et al. (2015b)'s samples at $z \sim 5.2$ and $z \sim 5.4$, so that small differences are expected between our measurement and a 'true' representation of cosmic variance in the same way as seen at lower redshifts.

At $z \sim 6.0$, our sample is smaller than all the ones used in Becker et al. (2015b) at lower redshift and the results should be interpreted with caution. We do find that some lines of sight are already transparent at $z \sim 6$, with roughly $\sim 40\%$ of our small sample having $\tau < 3$. In particular, three lines of sight display large transmission peaks as shown in Figure V.4.2. Quasars J2325–5229, J0143–5545 and J1148+0702 are located at $z = 6.23$, $z = 6.25$ and $z = 6.339$ respectively, and display the largest peaks of Lyman-α emission at $z = 5.9604$, $z = 5.9660$ and $z = 6.0914$ respectively. The first two spectra were taken on the EFOSC spectrograph while the third was obtained with the MMT instrument. Additional lines of sight are needed to establish whether such features are rare. In the case of the EFOSC spectra,

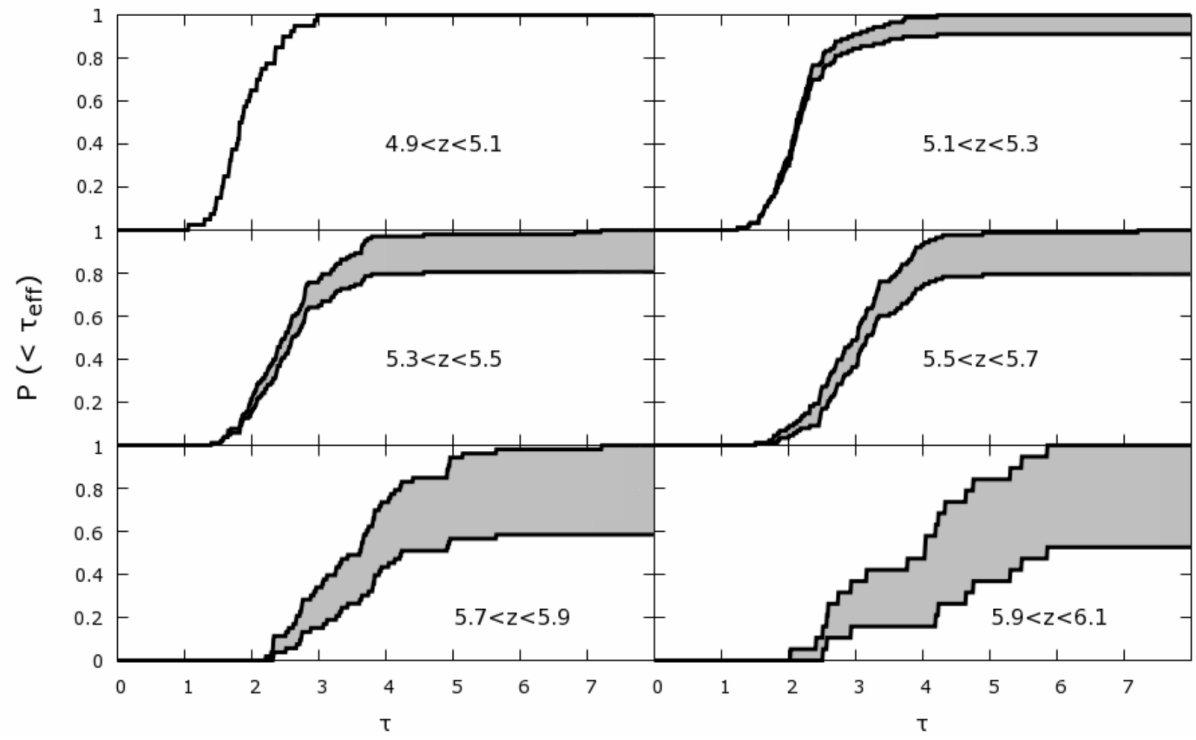

Figure V.8 – New results, plotting the most optimistic and most pessimistic contours based on the intrinsic values values of non-detections. The leftmost contour corresponds to non-detections have real values of twice the average error (as in Fig V.7) while the rightmost contour assumes non-detections are intrinsically maximally opaque (see text).

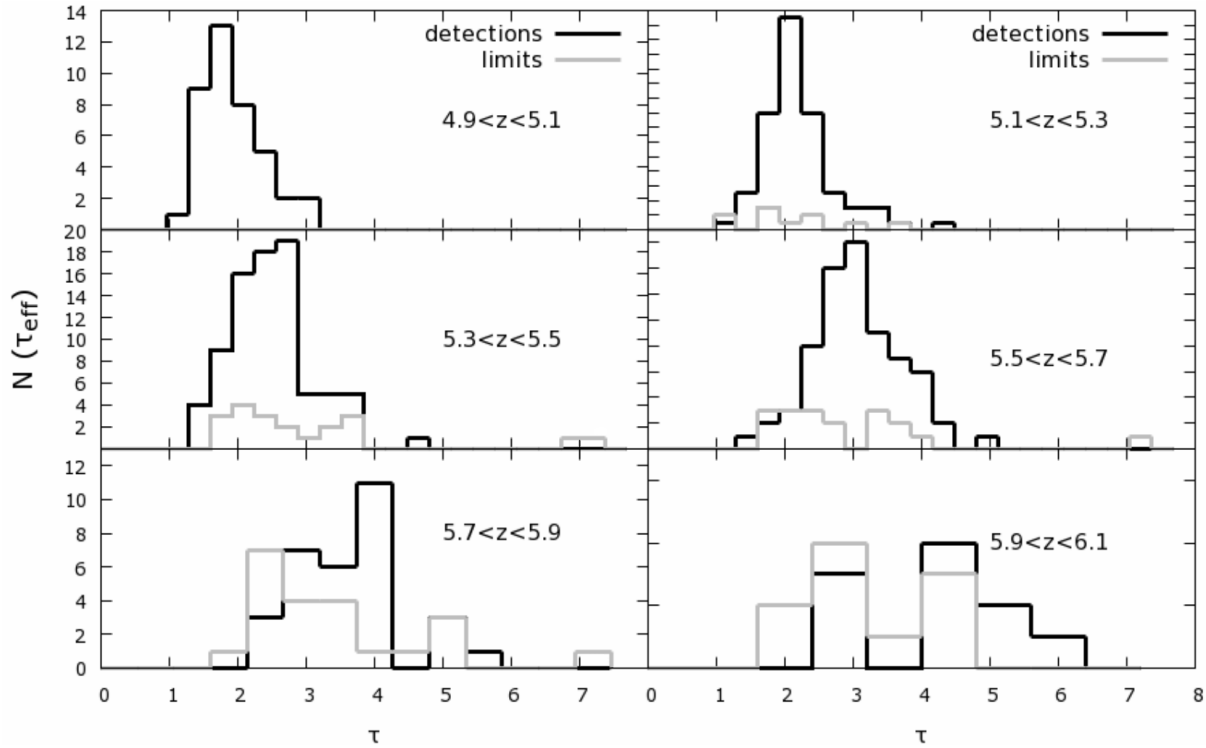

Figure V.9 – Differential distributions of opacity in 50 cMpc h^{-1} bins, showing detections in black and lower limits in gray.

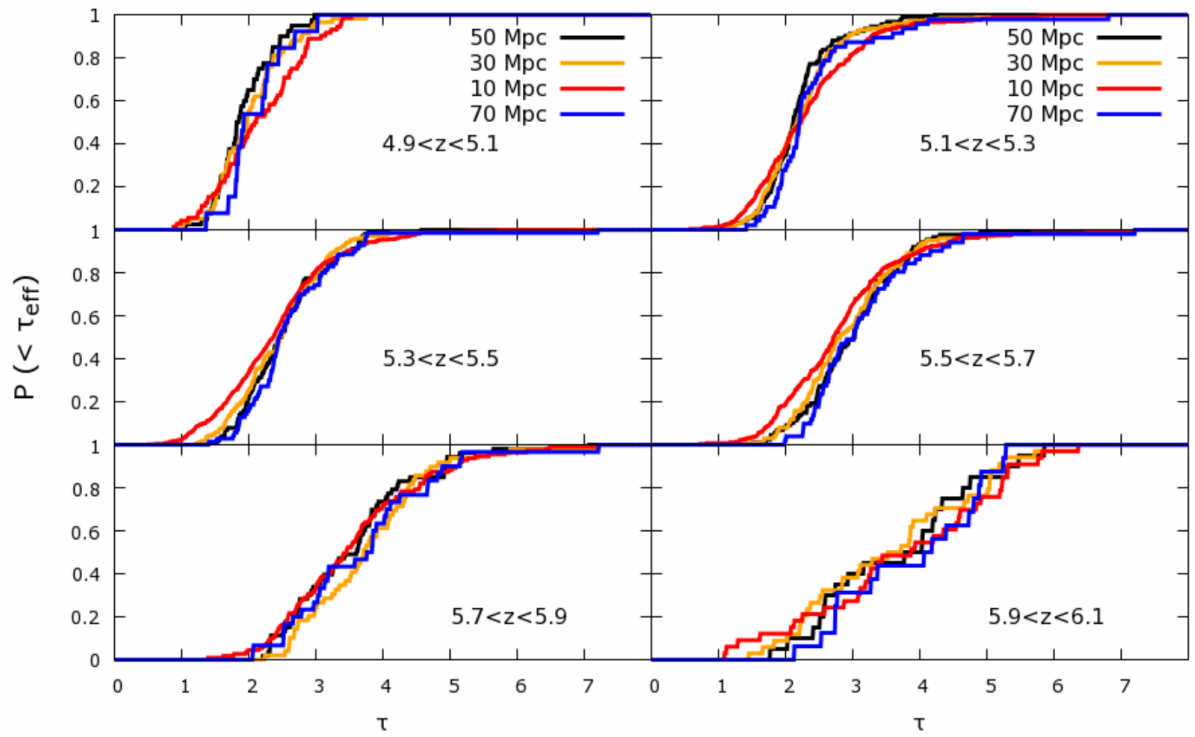

Figure V.10 – The effect of varying the size of the window over which Lyman-α transmission is measured. No significant effect is seen between windows of 30, 50 and 70 cMpc h^{-1} at any redshift, suggesting that fluctuations occur on even larger scales. Binning the data in 10 cMpc h^{-1} windows strongly affects the distribution, in particular at lower redshifts where it results in larger scatter.

more careful reduction might be needed to rule out the possibility of contamination. We also note that the peaks in J2325–5229 and J1148+0702 are located at rest frame wavelengths of $\lambda = 1172$Åand $\lambda = 1175$ Å respectively. This is within 6Å of the formal end of the proximity zone we adopt in this work. Better spectra of these objects are required to ensure the peaks are not related to particularly long and sporadic proximity zones.

In Figure V.9 we plot the differential distribution function of Lyman-α opacity across redshift. We distinguish between detections and lower limits using separate histogram colours. The distributions are clearly non-gaussian, with peak (most probable) values increasing linearly with redshift. The tail of opaque lines of sight at $z \sim 5.2$ is clearly visible and appears smooth and well sampled.

The effect of varying the size of the integration window is shown in Figure V.10. Although the effect is subtle, decreasing the window size tends to broaden the distribution, as expected. This is a natural consequence of cosmic variance. The broadening is particularly pronounced when the window size is decreased beneath 30 cMpc h^{-1}. The redshift range $5.1 < z < 5.7$ is more clearly affected, possibly because these redshifts are better sampled. We find no statistically significant difference between bins of 50 and 70 cMpc h^{-1} at any redshift.

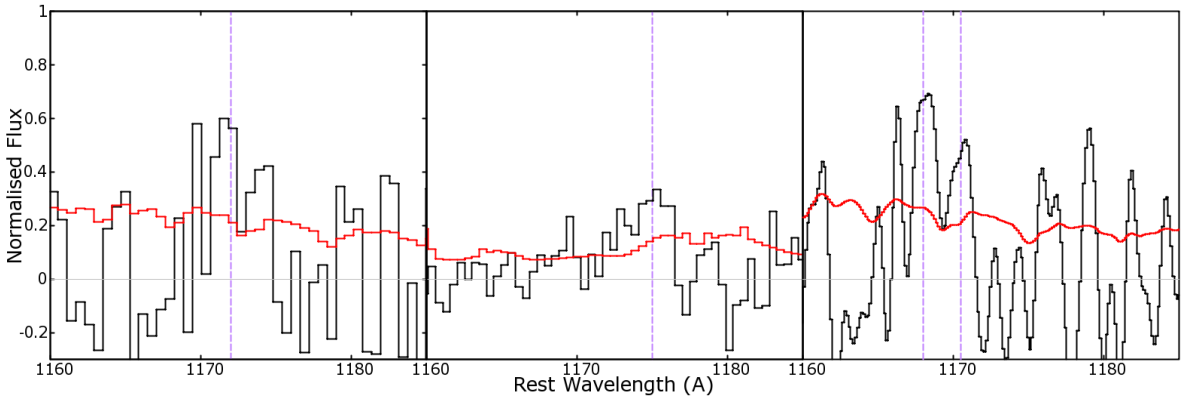

Figure V.11 – Transmission peaks detected above $z > 5.90$ in three quasar lines of sight. Spectra have been shifted to the rest frame. From left to right: EFOSC spectrum of J2325–5229, MMT spectrum of J1148+0702, and GMOS spectrum of J0143–5545. The spectra have been normalised.

V.4.3 Effect of data quality

We select sub-samples from the quasar catalog based on the signal to noise ratios. A 'GOLD' sample is chosen with SNR≥ 11.2 to match the SNR of the worst spectrum used in McGreer et al. (2015). This yields a sample of 35 high-quality spectra, more than halving the sample. Similarly, we construct a 'SILVER' sample by applying a less stringent cut of SNR≥ 5.3, matching the data quality from Eilers et al. (2017). This yields 51 lines of sight. We note that the sample of 33 quasars used by Becker et al. (2015b), the largest one so far, includes measurements from quasars with SNR much lower than these thresholds re-used from Fan et al. (2006) – down to SNR$= 3.2$. The results results are shown in Figure V.12. There is no significant trend with data quality at $z < 5.5$, where the distribution is well-sampled even by the small GOLD sample. A disparity appears at $z > 5.5$.

The differences between the full, SILVER and GOLD sample at $5.5 < z < 5.9$ are due to a combination of the plotting method, limited sensitivity, and small sample size. Indeed, we show here the 'optimistic' CDFs, meaning that non-detections in poor quality spectra are plotted at a fixed value of twice the average error. At $5.5 < z < 5.9$, the GOLD and SILVER samples CDFs fall within the bounds of the 'optimistic' and 'pessimistic' CDFs from Figure V.8. In other words, the GOLD sample distribution cannot be excluded at these redshifts. It should be noted that the data with SNR< 3.7 increasingly lacks the sensitivity to measure opacities larger than $\tau \gtrsim 4.0$ due to being too shallow. This explains the 'saturation' of the 'optimistic' bound around those opacities; the 'real' CDF is likely located closer to the SILVER sample.

At $5.9 < z < 6.1$, the difference is attributable to three lines of sight with detections of transmission, two of which are part of the SILVER sample. All other 'bins' plotted at low τ are lower limits, as can be seen in Figure V.8. In this case, the distribution inferred from the GOLD sample alone can be ruled out, as it falls outside the permitted bounds. In other words, three lines of sight revealed transmission despite inferior data quality, and the disagreements are therefore due to small sample size.

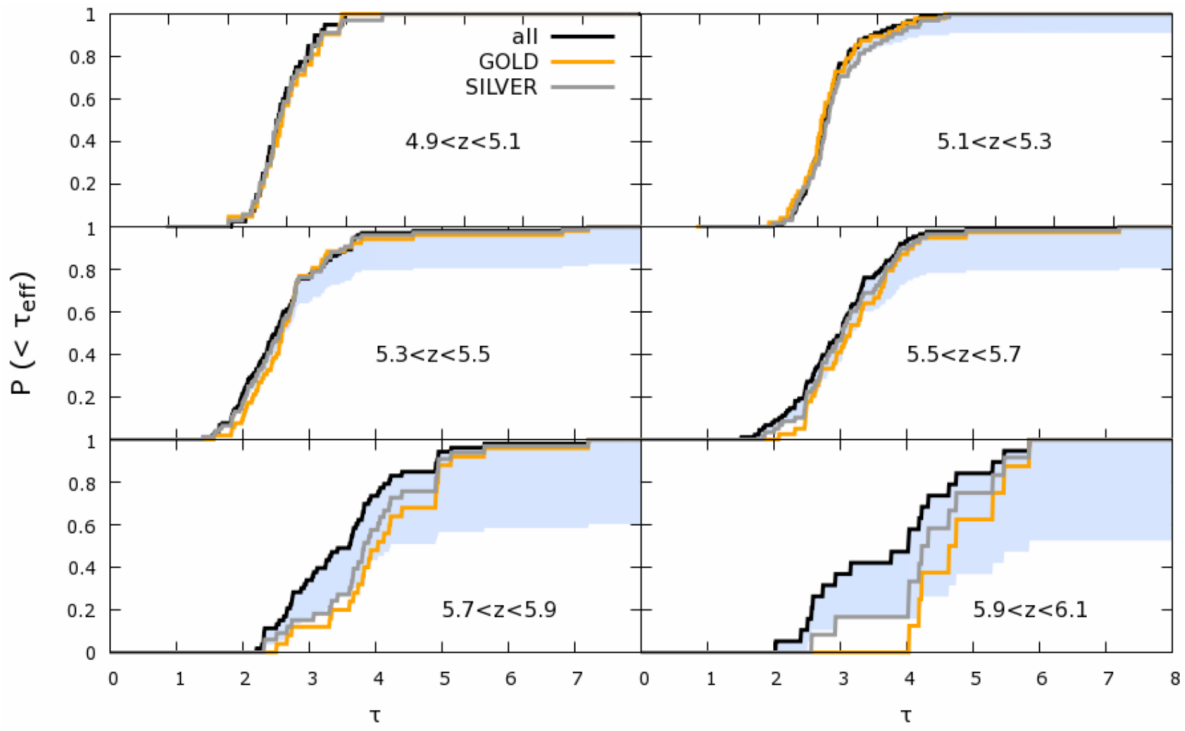

Figure V.12 – Cumulative distribution functions of Lyman-α opacity across redshift, computed using the full sample (black), the SILVER sub-sample of 51 objects matching the quality of data used in Eilers et al. (2017) (SNR$\geq$ 5.3), and the GOLD sample of 35 objects which match the quality from McGreer et al. (2015) (SNR$\geq$ 11.2). Shaded blue area shows the 'optimistic' and 'pessimistic' bounds presented in Fig. V.8. At $z < 5.5$, the distributions are well resolved by all samples and the distributions therefore agree. At $5.5 < z < 5.9$, the GOLD distribution lies between the bounds shown in Figure V.8. At $z > 5.9$, the difference is attributable to the small sample size of the sub-samples (see text).

z	Chardin et al. (2017b)	Keating et al. (2017)	Sherwood
5.0	N/A	1.459	0.569
5.2	N/A	1.391	0.541
5.4	0.814	1.397	0.525
5.6	0.731	0.898	0.514
5.8	0.564	0.733	0.474
6.0	N/A	0.368	0.335

Table V.4 – Emissivity rescaling factors (α) used to tune the simulations discussed here. The factors are chosen to match the observed flux F following $F = \langle \exp(-\alpha \cdot \tau_{\rm los}) \rangle$ (see text).

V.5 Comparison with models

In Figure V.13, we compare the updated CDFs of Lyman-α transmission in 50 cMpc h^{-1} windows with three simulations. Here, we briefly describe these simulations and outline their most relevant features.

When comparing predictions from these numerical simulations to observational results, it is important to keep in mind a few caveats. Firstly, all of the following numerical models explicitly re-scale the ionising emissivities to match the mean effective optical depth of observations, $\langle F \rangle_{\rm obs}$, by freely choosing a parameter A such that:

$$\langle e^{-A\tau_i} \rangle = \langle F \rangle_{\rm obs}. \tag{V.2}$$

At low redshift ($z \lesssim 5.7$), this rescaling is small and is somewhat justified by the difficulty of self-consistently generating the ionising UVB. In the following analysis, all simulation snapshots have been rescaled in this way to the corresponding 'optimisitic' (leftmost) CDF contour.

These simulations all use the UVB prescription of Haardt & Madau (2012) (HM12) as a starting point, which incorporates the best available estimates of the nature of ionising sources, the ratio of galaxy to AGN contributions, the escape fraction, the spectral shape of sources, and other factors. The precise values of these parameters are not known, and vary across redshift. The HM12 emissivity is therefore used as a best guess, and the final re-scaling of the emissivity reflects the known inaccuracy in the models. The rescaling is therefore simulation– and redshift–dependent, but also resolution–dependent since the small-scale recombination and self-shielding effects necessary to calculate the UVB self-consistently are currently beyond the reach of numerical simulations. In Table V.4 we list the values used to rescale in optical depths in each case (α), chosen to match the observed flux F following $F = \langle \exp(-\alpha \cdot \tau_{\rm los}) \rangle$. At high redshift ($z \gtrsim 5.9$) this rescaling procedure becomes less valid because the correction is not small. For instance, the mean flux values of the Sherwood simulation used here had to be re-scaled by a factor > 10 at $z = 6$. This reflects the fact that the HM12 background is a bad representation of the ionising emissivity at high redshift, and ideally should not be used.

Secondly, the timing of reionisation is a free parameter in most of the following models. The Sherwood simulation, and the HM12 ionising background, were designed to fit Lyman-α transmission at lower redshifts than explored here. Because the time evolution of Lyman-α transmission is slower at $z < 4.9$, the successful predictions of these cosmological simulations are usually robust to shifts up to $\Delta z_{\rm Re} \sim 0.2$ over the redshift range considered here (see e.g. Chardin et al. 2017a).

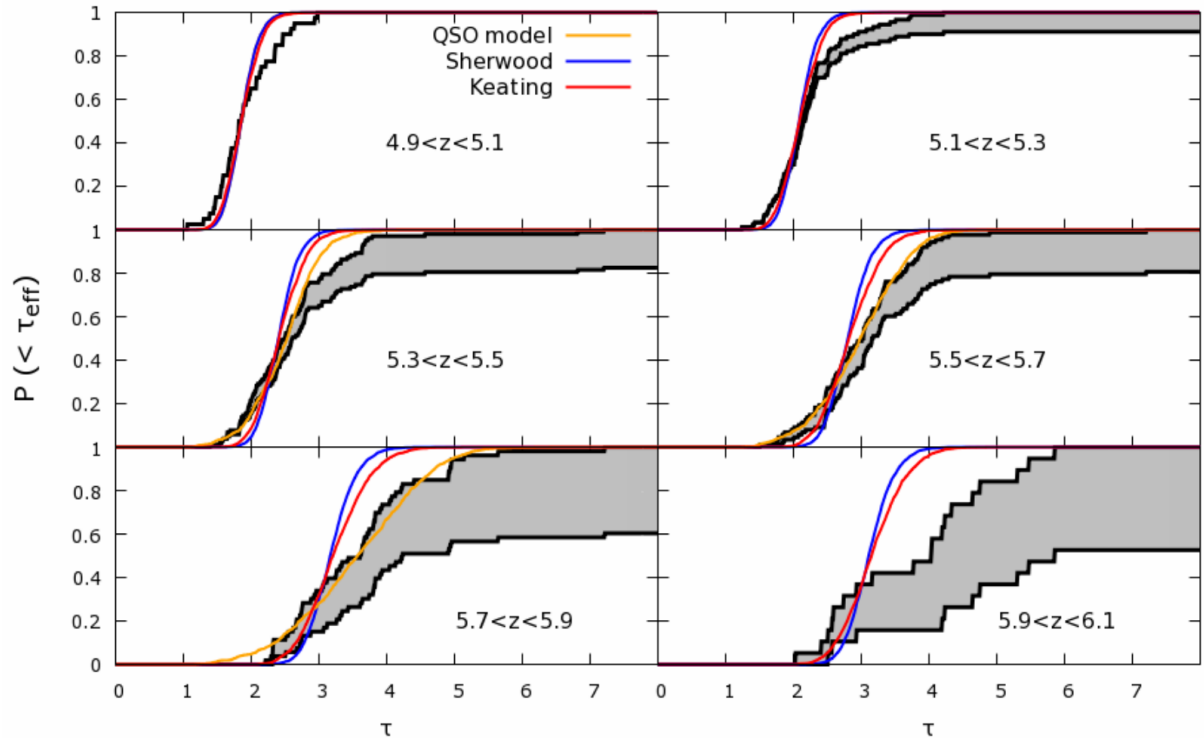

Figure V.13 – Comparison of the measured Lyman-α PDFs at $4.9 < z < 6.1$ with outputs from a range of numerical simulations. The shaded areas correspond to treating non-detections optimistically or pessimistically as in Fig V.8.

V.5.1 The Sherwood Simulation

The Sherwood simulation suite (Bolton et al., 2017) was designed to reproduce the Lyman-α forest over $2 < z < 5$ and test its sensitivity to a range of model parameters such as galactic and AGN-driven outflows, thermal histories, and cold/warm dark matter. With gas particle masses of $M_{\rm gas} = 9.97 \cdot 10^4 M_\odot$ and 2×2048^3 particles for a box size of 40 cMpc h^{-1}, the Sherwood simulation used here possesses higher mass resolution than other larger-scale cosmological simulations such as Illustris and EAGLE. The simulation is run using the hydrodynamics code P-GADGET 3 (Springel, 2005) and used $\sim$ 2680000 core-hours of computing time. Bolton et al. (2017) compare simulated Lyman-α lines of sight with data from Fan et al. (2006), finding remarkable agreement over $2.5 < z < 5$ with a slight lack of opaque systems at $z = 2.0$ – possibly related to the inability of the simulation to resolve self-shielding. Because it was designed to match observations at redshifts less than $z = 5$, the Sherwood simulation uses a uniform UVB with a shape given by the HM12 model, in a scenario where reionisation is driven mostly by galaxies and with no radiative transfer. The interrelations of neutral fraction, temperature, photon mean free path, and density are therefore not taken into account, and the simulations are expected to fail in the hydrogen reionisation regime where the UVB is known to be inhomogeneous.

Here we compare the 40 cMpc h^{-1} Sherwood simulation box with the Lyman-α CDF over $4.9 < 6.1$. Each measurement of $\tau_{\rm eff}$ requires stitching together two simulated lines of sight; 2500 total values of $\tau_{\rm eff}$ are obtained. The model predictably falls short of reproducing the variety of line-of-sight opacities at $5.3 < z < 6.1$. It is perhaps more surprising that the

uniform UVB model also fails dramatically at $5.1 < z < 5.3$, where line-of-sight variation is slightly more pronounced than previously reported. Percolation of UV photons is predicted to have occurred long before $z = 5.2$ by most models.

V.5.2 Radiative transfer simulation

In addition to the homogeneous-UVB Sherwood simulations, we compare our results to the full cosmological radiative transfer simulations of Keating et al. (2017). Following D'Aloisio et al. (2015), these simulations test the effect of regions ionised at different redshifts on the spatial variations in the temperature-density relation. The temperature dependence of recombinations rates could then lead to increased fluctuations in Lyman-α opacity. Unlike previous models, Keating et al. (2017)'s simulations use an extended and self–consistent reionisation history to boost the IGM temperature. The injected energy and history are chosen to match the temperature and photo-ionisation rates of the IGM at $z \lesssim 5.0$ measured using the Lyman-α forest.

This difference to previous models turns out to be important, as Keating et al. (2017) find their more realistic choices of reionisation heating do not reproduce the large Lyman-α opacity fluctuations previously reported. Choosing a higher ionisation energy to match the IGM heating in D'Aloisio et al. (2015) they find the fluctuations are still not large enough, and moreover the produced lines of sight are in tension with low redshift Lyman-α forest data as well as transmission peak statistics at high redshift.

The simulations are also run in P-GADGET 3 but snapshots of the resulting density field are then post-processed to include the effects of temperature and ionisation using mono-frequency radiative transfer. Here we use a 40 cMpc h^{-1} box with 512^3 mass particles, which should better capture large-scale variations of opacity than the higher resolution 20 cMpc h^{-1} box. Keating et al. (2017) compare their Lyman-α PDFs to data over $4.9 < z < 5.9$ and find no trend with box size. Here we stitch together two simulated lines of sight for each of the 2500 total measurements of $\tau_{\rm eff}$. We find that this model, while performing better than the Sherwood simulation, does not provide a satisfactory fit to the Lyman-α CDFs. Although including radiative transfer does increase line-of-sight variance, the effect is too small by at least factor of two. This is in agreement with the results of Keating et al. (2017).

V.5.3 Rare sources simulation

Finally, we compare the results to predictions of a quasar-driven reionisation toy model from Chardin et al. (2015). These models include a density field produced by the hydrodynamics code RAMSES onto which radiative transfer is added in post-processing with the ATON code. They found the fluctuations in the UV produced by galaxies in the redshift range $5 < z < 6$ to be on scales too small (< 50 cMpc) to account for the spread in Lyman-α opacity. A much better fit was found when rare, bright ionising sources were added to the simulation with a carefully chosen spatial density. As discussed in the Introduction, such sources could be (faint) AGN.

The simulation boxes of Chardin et al. (2015) are unfortunately only 20 cMcp h^{-1} in size, but this is not a problem as the line of sight variance is mostly driven by the presence or absence of a strong ionising source nearby. The simulations use 512^3 mass particles. Because of the smaller simulation volume, we have to stitch three simulated lines of sight together for each

measurement of $\tau_{\rm eff}$. This results in 1020 measurements a smaller number than in the previous cosmological simulations.

We find that the rare sources simulations aptly reproduce the CDF at $5.5 < z < 5.9$, and are the only model to do so out of the ones we tested. We note however that the line of sight variance seems to disappear very quickly once percolation has happened, resulting in the model under-estimating the opacity variance at $5.3 < z < 5.5$. Unfortunately no snapshots were available at $5.1 < z < 5.3$, but it is unlikely that those would show a sufficient amount of opaque lines of sight since those are already gone from the simulations at $z \sim 5.4$. However, we note that there might be room for reionisation to occur later within this model, which could potentially ease the conflict. Chardin et al. (2015) also note that their simulations are in better agreement with the Lyman-α CDFs calculated with smaller window sizes of $l = 10$ cMpc h^{-1}. This is still the case with the updated measurements, as can readily be seen by comparing Figures V.10 and V.13.

There is also tension between this model and observations for low-opacity lines of sight at $z \sim 5.8$: in a QSO-driven scenario, low opacity regions arise purely due to proximity to a source AGN and are therefore not expected to disappear completely at high redshift, as long as the QSOs are active. This appears to be in contrast with observations which do report that transparent lines of sight go missing at $z > 5.7$. However, this problem is mitigated by

1. the smaller observed sample size at these redshifts means the discovery of even 1 transparent line of sight could ease the conflict, and

2. while the ionising emissivity is already tuned to observations, and it might be sensible to rescale the predicted opacity by a larger factor to improve the agreement with data.

We note that the volume density of rare bright sources in this model was explicitly chosen to reproduce the Lyman-α CDFs of Becker et al. (2015b) which use 'optimistic' measurements i.e. following our leftmost CDF contour. It is conceivable that the model could be modified to reproduce a CDF closer to the mid-point of our contours.

V.6 Conclusions

We have assembled a sample of 92 optical spectra of quasars with $z_{\rm source} \geq 5.7$ in order to measure the distribution of IGM Lyman-α transmission over $5 < z < 6$. These objects consist of 52 SDSS-discovered quasars (23 of which have not been re-observed), 18 quasars from DES-VHS, 9 from the SHELLQs survey, 13 from online telescope archives, 23 from previous similar studies and 4 new spectra. The data originates from a total of 9 different optical spectrographs and has been collected over the course of the last 11 years. We use this unprecedented sample of high-z quasars to improve the measurements of residual Lyman-α transmission of Fan et al. (2006) and Becker et al. (2015b). The large variance in opacity among lines of sight has previously been shown to be incompatible with a uniform UVB. At the same time, unexpectedly large longitudinal correlations in opacity of up to 110 cMpc h^{-1} at $z > 5.5$ mean that only a dramatic increase in number of background sources – and not a finer sampling – can aptly quantify cosmic variance (Becker et al., 2015b).

Our new constraints on the Lyman-α cumulative distribution function (CDF) are shown in Fig V.8. In past work, upper limits on transmitted flux were dealt with in a "maximally optimistic" way whereby a value of twice the error was adopted, i.e. the real flux was assumed

to be just below the detection threshold. Using this technique, we demonstrate that our results are compatible with previous studies within bootstrap sampling. In order to deal with the disparate data quality in our sample and present limits in a more transparent way, we introduce a second bound on the CDF which is "maximally pessimistic", i.e. non-detections are taken to mean that $\tau \rightarrow \infty$. This allows us to incorporate moderately deep data while being confident that the 'true' CDF lies in-between these two bounds; in other words our results present the region permitted by current data.

We test the effect of altering the comoving size of the binning window, and find only subtle effects on the distribution between the $l = 30, 50$ and 70 cMpc h^{-1} cases at any redshift. A binning with $l = 10$ cMpc h^{-1} makes the distribution of effective optical depth significantly broader at $z < 5.7$.

We also vary the length of the excluded "proximity region" which is affected by the source quasar itself, finding no effect at any redshift on the statistical distribution of transmitted flux as long as $\lambda_{\mathrm{end,\,prox}} < 1180$ Å is adopted. The traditionally used value of $\lambda_{\mathrm{end,\,prox}} = 1176$ Å is thus valid and we don't expect significant contamination from the quasar proximity zones.

We compare our final results with outputs from three different published numerical models: the constant-UVB Sherwood simulation (Bolton et al., 2017), the radiative transfer post-processed simulation of Keating et al. (2017) which model temperature fluctuations arising from differences in the timing of reionisation, and a model including rare, bright sources from Chardin et al. (2015). We tune the UV emissivities of all simulations at all redshifts to match the average flux measured from our sample in each redshift range. We find that the evidence strongly rules out the constant UVB model and the radiation transfer model in their current forms. The rare sources model is marginally consistent with data at $z > 5.7$ but not at $5.5 < z < 5.7$. However, we note that these models contain free parameters which have been chosen to match older observations, and more work is required in order to rule out or confirm different reionisation scenarios.

A sample of the size presented here opens up new avenues for exploring the epoch of reionisation. As an example we present results from the recent Chardin et al. (2017a) paper where the occurrence rates of individual Lyman-α transmission spikes in the spectrum of J1120+0641 are found to be consistent with the post-processed radiation transfer simulations of Chardin et al. (2015) in which reionisation is driven mostly by galaxies. The predicted distribution of spike strengths from that model is stable to changes in the UVB emissivity and timing of reionisation. This is promising for future studies which will incorporate more lines of sight and more numerical models.

Another potential future application of this sample is the characterisation of quasar proximity zones and their evolution across redshift. Proximity zones have been reported to shrink over $5.7 < z < 6.4$ (Bolton & Haehnelt, 2007a; Calverley et al., 2011; Keating et al., 2015; Eilers et al., 2017), a trend we can confirm and push to higher redshifts. Additionally, quasars with low enough redshifts or X-Shooter infrared data could be used to measure the spectral emission properties of high-z quasars with the largest sample yet. Evolution in the C IV or Si IV emission lines or the quasar continuum parameters could signal important changes in the underlying accretion processes, while an evolution of the Lyman-α emission line could be linked to the emergence of Gunn-Peterson damping wings during reionisation. Finally, many of the objects presented here have only been targeted to moderate depth with modest instruments, and provide good targets for ESI, HIRES or X-Shooter. This sample offers a great starting point for a potential large-scale spectroscopic campaign of early AGN.

Acknowledgements

This research has made use of the Keck Observatory Archive (KOA), which is operated by the W. M. Keck Observatory and the NASA Exoplanet Science Institute (NExScI), under contract with the National Aeronautics and Space Administration. Based on data obtained from the ESO Science Archive Facility. This paper used data obtained with the MODS spectrographs built with funding from NSF grant AST-9987045 and the NSF Telescope System Instrumentation Program (TSIP), with additional funds from the Ohio Board of Regents and the Ohio State University Office of Research.

V.7 Extra Figures

V.7.1 Mosaic of quasar catalog

Here we plot all the $z > 5.7$ quasar spectra used in this work in a common format. The spectra are normalised by dividing the flux by a best fit power-law as described in Section V.3. Wavelengths are divided by $z_{\text{source}} + 1$ to bring the spectra into the rest frame. The y axis is calibrated so that it spans the range $0 \rightarrow 5\times$continuum for each quasar. We do not bin the spectra in order to reflect the diversity of data qualities present in the sample. Error arrays are shown in red.

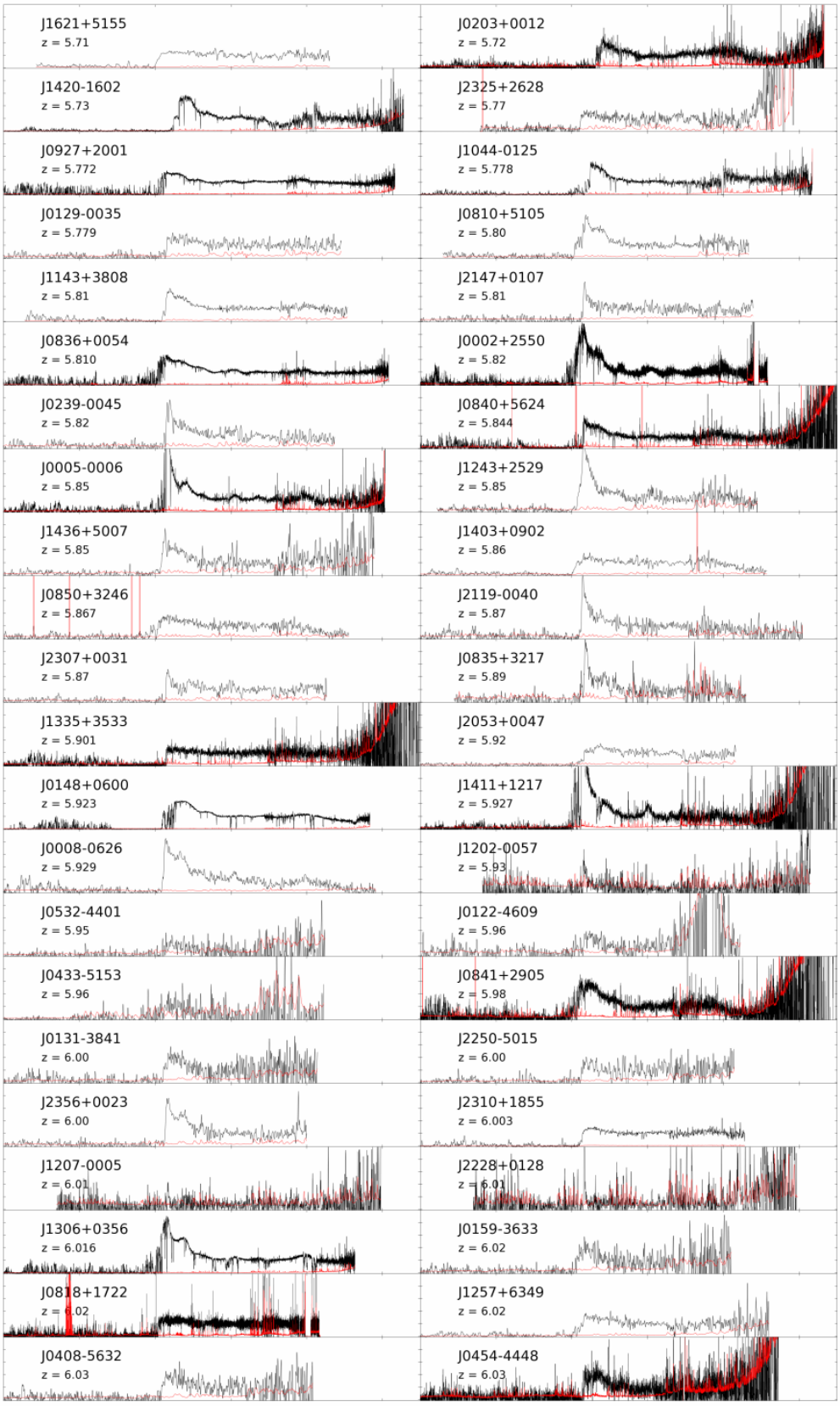

Figure V.14 – First half of the quasar catalog. The origin of each spectrum and the instruments used are listed in Table V.2. Wavelength runs from 1000 to 1550 Å and the fluxes have been normalised by dividing by the best-fit power law to the continuum.

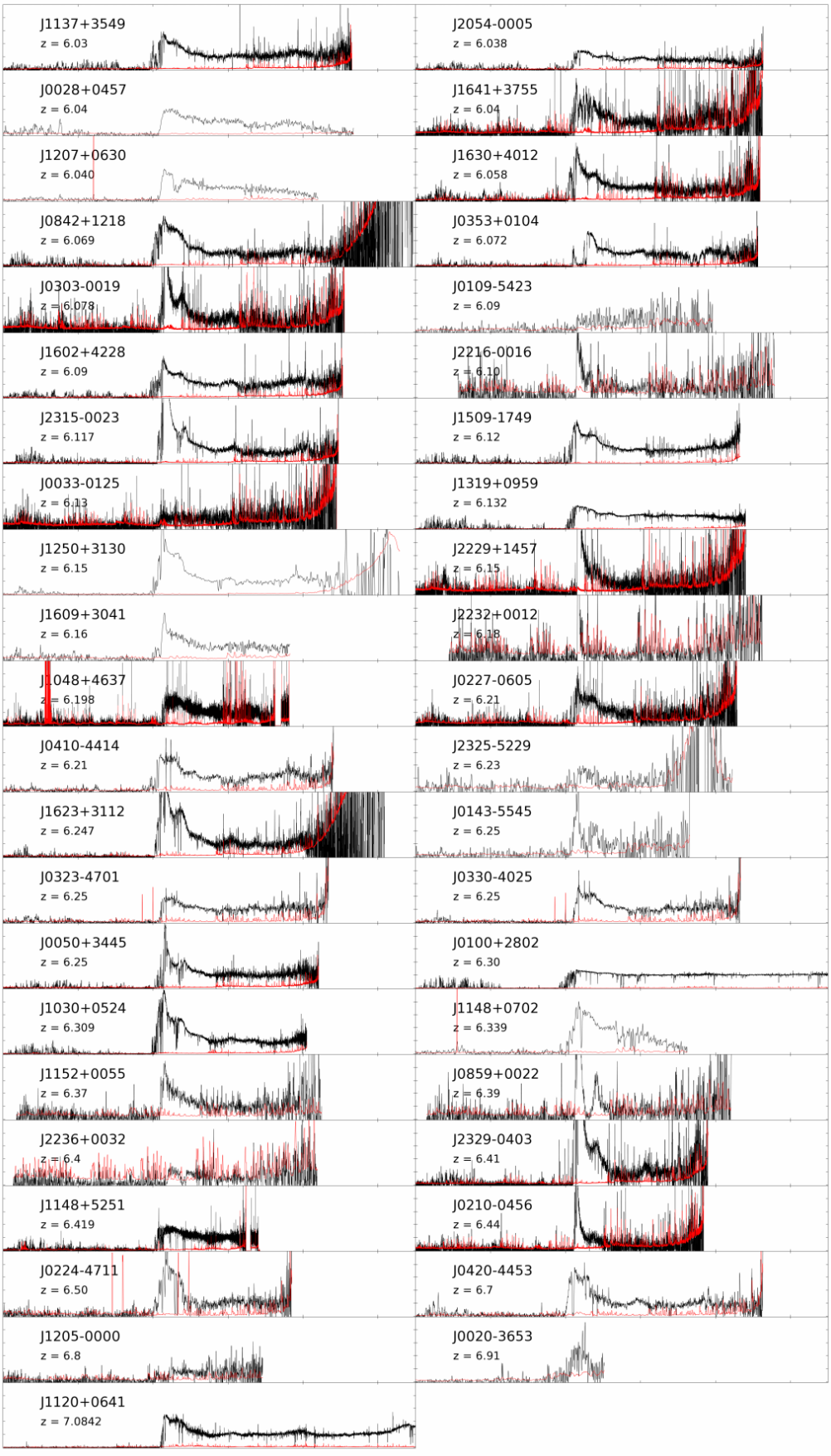

Figure V.15 – Second half of the quasar catalog. Data is as in Fig V.14.

VI Conclusions

VI.1 Summary

In recent years rapid progress has been made in uncovering the events of the first billion years on both the observational and theoretical fronts. Of particular interest are the processes of galaxy formation and reionisation. The properties of the IGM, metal enrichment, and the UVB among other cosmic properties have been explored at increasingly early times and found to be evolving. However, despite numerous observations, we have so far only studied the tail end of reionisation and the clues gathered on the process are ambiguous. Lyman-α emitters go missing at $z > 6$, but many processes unrelated to the IGM influence their visibility. High-redshift quasars have yielded contradictory constraints on the hydrogen neutral fraction based on the Gunn-Peterson damping wing and dark gap statistics. Metal enrichment is a puzzle, as the increase in O I occurrence at high redshift seems to require a weakening of the UVB while models also under-produce C IV systems, which require a stronger UVB. The relevance of galaxy formation processes to reionisation has even been called into question with the emergence of models which include a significant contribution from quasars to the reionisation photon budget.

The objective of this thesis has been to provide new observational constraints based on high-redshift quasar spectra to help resolve these puzzles. The main achievements of this thesis are summarized as follows:

- We have found analogues to the $z = 7.085$ quasar ULAS J1120+0641 (J1120) at lower redshift in terms of their C IV and Lyman-α emission line properties (Chapter III). These objects were found in the SDSS DR12 catalog at $2.4 < z < 3.4$. The existence of these analogues casts doubt on whether the strength of the Lyman-α emission line of J1120 is necessarily an indicator of damping wing absorption due to neutral gas along the quasar line of sight.

- Moreover, we have established that the Lyman-α emission line strength in J1120 is not anomalous among a population of lower-redshift quasars selected to match its C IV prop-

erties. This is in contrast with previous studies which constructed quasar spectra composites to establish that J1120's Lyman-α line was absorbed. We showed that the small mismatch in C IV properties, of about 3 Å in the rest frame, between that composite and J1120's C IV line may be sufficient to explain its weak Lyman-α line. This confirms that the relation between Lyman-α and C IV emission in quasars is complicated.

- Using a deep (30h) X-Shooter spectrum of J1120, we measured abundances of multiple metal ions at the higher redshifts yet (Chapter IV). We detected a single C IV system at $z = 6.51$, the highest redshift for an intervening absorber so far. We measured the C IV mass density at $z > 6.2$ and found it to be consistent with either non-evolution or decrease from $z = 5.5$. We also find one C II absorber at $z = 6.40$, consistent with a non-evolving incidence rate from $z = 6.0$.

- We discovered four weak ($W < 0.3$ Å) Mg II systems at $z > 5.95$ along the line of sight to J1120, which is significantly more than expected from a power-law extrapolation of the number density of stronger systems. A similar over-abundance of weak Mg II systems has been reported at $z < 2.5$, possibly associated with metal-enriched outflows. The processes producing weak Mg II systems may therefore have been established very early and remained relatively stable over time.

- We found evidence of partial covering of J1120's broad line region by an absorber situated -1100 km s^{-1} from the source. The peculiar absorption profile of C IV and N V in this absorber are consistent with a covering fraction of $\sim 40\%$. We cannot rule out that the unusual profiles are instead caused by multiple narrow, cold ($T < 5000$K) absorbers located very close to each other. If the source is indeed partially covered, it is likely that a Lyman limit system is present near J1120 even though some Lyman-α transmission is present. This would make the quasar's proximity zone appear artificially short without the need for significantly neutral IGM.

- I have assembled a large sample of 92 optical quasar spectra in the redshift range $5.7 \leq z_{\text{source}} \leq 7.0$ (Chapter V). This in an increase by a factor of 2.8 compared to the previous largest sample of such objects. We brought together data taken by the SDSS, DES–VHS and SHELLQs surveys, as well as spectra taken for previous studies of high redshift Lyman-α transmission, archival data, and four new spectra. This sample opens up multiple avenues for future studies of high-z quasars.

- We introduced a new method to treat upper limits on Lyman-α flux in the measurement of the line-of-sight Lyman-α transmission in quasar spectra. Combined with the traditional methods of quantifying the Lyman-α opacity distribution, this allows us to put 'most optimistic' and 'most pessimistic' bounds on the distribution of Lyman-α flux.

- Using this sample, we measured new probability distributions of IGM Lyman-α opacity at $4.9 < z < 6.1$. These measurements are currently the best representation of cosmic variance in IGM opacity at these redshifts. We compare the measured distributions with three numerical simulations and confirm that IGM models with a spatially homogeneous UVB are strongly excluded. A reionisation model with post-processed radiative transfer deign to test the effect of temperature fluctuations from Keating et al (in prep) eases the tension but falls short of reproducing the observed scatter in Lyman-α opacity by a

factor of a few. A model where the UVB is dominated by rare sources from Chardin et al. (2015) provides the best fit to data out the models we tested, but still lack sufficient opaque lines of sight at $z < 5.5$. It also struggles at other redshifts if the measurements are treated as limits.

- We have investigated the effect of window size on the inferred Lyman-α transmission distribution, which is usually averaged over 50 cMpc h^{-1} regions. We find only subtle differences among the distributions using windows of 30, 50 or 70 cMpc h^{-1} at any redshift, although a stronger effect is seen with bins of 10 cMpc h^{-1}, particularly at $z < 5.7$, where the small bin size results in a broader distribution.

VI.2 Outlook

Several opportunities for future projects have arisen over the course of this thesis. Some of these are ongoing research, while some are potential extensions of the work presented here. In this section I briefly outline upcoming research.

VI.2.1 Other applications of the $z > 5.7$ quasar sample

While gathering data in order to measure the Lyman-α mean flux distribution towards high-redshift quasars, we have assembled a large sample of 92 objects with $z_{\rm source} > 5.7$, including 120 total spectra (some objects were observed more than once). These spectra have been normalised, stored and (when possible) reduced in a consistent way, and thus offer a unique opportunity to investigate the spectral properties of high redshift quasars. Some future avenues include:

- The emission properties of quasars can be studied across a wide range in redshift. Currently the best sample of optical quasar spectra is the SDSS DR12 sample (Pâris et al., 2017) containing 297 301 quasars spanning mostly the range $0.0 < z < 4.5$ (184 101 objects at $z > 2.15$). A comparison with our sample could reveal trends at very high redshift in, e.g., quasar continuum properties, with potential implications for the dust obscuration of early AGN. Additionally, a comparison of the properties of, and links between, broad emission lines would confirm or challenge the assumption that the intrinsic properties of AGN do not evolve with redshift.

- The evolution of proximity zone sizes across redshift can be probed with the largest sample yet. Previous studies have found a large scatter in the proximity zone lengths of individual quasars, requiring more data to quantify these trends. This would require improved redshifts for some of the quasars in the sample.

- With multiple exposures of the same quasars with the same spectrograph distributed across multiple years it becomes possible to check early AGN for variability. At $z < 0.7$ is it estimated that roughly $4 - 12\%$ of quasars display strongly variable continuum emission in the wavelengths covered by this sample, and $1 - 2\%$ of those show appearing or disappearing broad emission lines over time-scales of 8–10 years (see MacLeod et al.

2016 and references therein). The identification of any variable AGN so early in the Universe would open investigations into the accretion processes at work in these objects, a crucial aspect of AGN feedback models.

- Multiple exposures of the same object with different spectrographs allow for a critical comparison of instrument biases in the context of high-redshift objects.

- Many of the spectra used here have only modest resolution or signal-to-noise ratios. This offers an ideal starting point for a large campaign to obtain spectra of high redshift quasars with sufficient resolution and depth to reveal additional extremely long Lyman-α troughs or other oddities.

VI.2.2 Characterising Lyman-α transmission spikes

Beyond $z \gtrsim 5.9$, Lyman-α transmission along quasar lines of sight occurs in discrete, often widely separated spikes. Measuring the transmitted flux over a redshift range by integrating the flux in those spikes alone, as has been done above and in previous studies (Fan et al., 2006; Becker et al., 2015b) does not reflect this property. More information can be gained from looking at the clustering, strength distribution, and occurrence statistics of transmission spikes (Gallerani et al., 2006, 2008; Feng et al., 2008).

The physical origin of transmission spikes varies between models: in a scenario where Lyman-α opacity scatter is driven by variations in the mean free path of ionising photons (Davies & Furlanetto, 2016), over-dense regions of the IGM will be more transparent to radiation as the mean free path is longer and the local ionizing emissivity is higher. On the contrary, if opacity scatter is driven by temperature differences resulting from differential timing of reionisation (D'Aloisio et al., 2015), it is the dense areas which will be transparent since the gas there will be recently ionized and therefore still relatively hot, with a low recombination rate (see also Xu et al. 2014). In a quasar-driven reionisation model, transmission spikes originate from moderately dense or under-dense region of the IGM which happen to lie close to an AGN.

There is therefore some hope of discriminating between models based on their varying predictions on Lyman-α spike statistics. For instance, in Chardin et al. (2017b) we note that the seven transmission spikes found the spectrum of the $z = 7$ quasar ULAS J1120+0641 are consistent with a model that includes a fluctuating UVB due to galaxies, but only if reionisation occurs sufficiently early (see figures 7 and 8 of Chardin et al. 2017b). More precisely, the model is completely consistent with J1120's residual transmission if the global opacity reaches $\tau = 8.0$ at $z = 6.5$, marginally consistent if this occurs at $z = 6.3$, and inconsistent if $\tau = 8.0$ is reached at $z \gtrsim 6.0$. Moreover, the strength distribution function of transmission spikes in Chardin et al. (2017b)'s models is independent from the timing of reionisation and re-scaling of the source emissivity. This opens the possibility of falsifying the models if the observed distribution of spike strengths does not follow this robust prediction, making transmission spike statistics an exciting new way to extract information from Lyman-α transmission at the highest redshifts.

VI.2.3 Searching for weak metals along quasar lines of sight

The work presented in Chapter 4 can be expanded upon in a variety of ways. A first step would be to combine our Mg II intervening system detections with those presented in a recent paper by Codoreanu et al. (2017). Probing multiple lines of sight, they find stronger Mg II systems as well as weaker ones. If the combined distribution lends itself to a fit by a double power law similar to the distribution seen at $z \sim 2$ this would provide important information about the morphology of the first metal-enriched regions in the Universe.

Another possibility would be to undertake a large survey of weak metals along quasar lines of sight at $z > 4.5$. Enough high-SNR X-Shooter spectra at $z_{\rm source} \gtrsim 5.0$ are now available to robustly probe the weakest C IV systems with $\log N_{\rm C\,IV} \lesssim 12.5$. The weak end of the column density distribution of C IV is poorly constrained at $z > 5.0$. Its determination ties in to galactic feedback processes by measuring the densities down to which the IGM has become enriched at early times. Since models struggle to reproduce the abundance of strong C IV systems, this might be key information for uncovering how carbon is distributed at those epochs. Any serendipitous co-occurrence of a Lyman-α transmission spike with metals would also have consequences for competing models of reionisation, some of which predict that low-opacity regions can only correspond to voids, rather than the high-density regions more likely to be enriched with metals.

VI.2.4 Follow-up of intervening absorbers with IFU spectroscopy

From an observational standpoint, the origin of high-z intervening metal systems is largely unknown. There have been some efforts to correlate metals absorbers with Lyman-α emitting galaxies, but these have been limited in scope and so far revealed few, if any, bona fide associations (Bouché et al., 2007; Steidel et al., 2010; Turner et al., 2014; Burchett et al., 2016; García et al., 2017). There is a clear interest in establishing this connection, since studies of metal-enriched gas have provided powerful insights into the inflow and outflow processes that regulate galaxy growth at $z \leq 3$. Until now, making this connection at higher redshifts has been observationally prohibitive due to the large time investment needed for imaging and follow-up spectroscopy of galaxies at $5 < z < 6$. An alternative approach, however, is to target quasar fields known to contain metal absorbers with IFU spectroscopy (and the MUSE instrument in particular). Bacon et al. (2015) have shown that MUSE detects large numbers of Lyman-α emitting galaxies (LAEs) in blank fields with relatively modest exposure times. Since the Lyman-α emission line provides a precise redshift, this approach offers an efficient means of probing galaxy-absorber connections.

Detecting any correlation between LAEs and intervening metal systems would immediately put constraints on the interplay of stellar formation and metal enrichment. The kinematics of the observed lines would yield further information about early feedback mechanisms. At the same time, a lack of correspondence could imply that metal lines are associated with galaxies that are intrinsically faint at least in terms of their Lyman-α emission, making them potentially exciting targets for *JWST* and E-ELT. It would also mean that LAEs are not, in general, the objects most responsible for metal enrichment at high redshift. Alternatively, the galaxies may not show Lyman-α in emission, which would carry implications for the magnitude of local effects such as the Lyman-α escape fraction.

VI.3 Closing Remarks

High redshift quasars are likely to remain a prominent tool for the study of the early Universe in the foreseeable future. Advancements in instrumentation and statistical techniques have provided spectra of higher quality and in greater number than ever before, and this progress will accelerate with the advent of 20m-class telescopes, adaptive optics, and new, more powerful spectrographs such as ESPRESSO on the VLT, ETC/HIRES and spectrographs on the GMT and TMT. Many new discoveries should therefore lie ahead as we use these incredible tools to probe the subtle interactions at work in the intergalactic medium during and after the epoch of reionisation.

Bibliography

Alam, S., Ata, M., Bailey, S., et al. 2016, ArXiv e-prints, arXiv:1607.03155
Alvarez, M. A., Busha, M., Abel, T., & Wechsler, R. H. 2009, ApJ, 703, L167
Amorín, R., Pérez-Montero, E., Contini, T., et al. 2015, A&A, 578, A105
Atek, H., Kunth, D., Schaerer, D., et al. 2014, A&A, 561, A89
Bacon, R., Brinchmann, J., Richard, J., et al. 2015, A&A, 575, A75
Bahcall, J. N., & Peebles, P. J. E. 1969, ApJ, 156, L7
Barai, P., Viel, M., Borgani, S., et al. 2013, MNRAS, 430, 3213
Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Barnes, L. A., Haehnelt, M. G., Tescari, E., & Viel, M. 2011, MNRAS, 416, 1723
Barnett, R., Warren, S. J., Becker, G. D., et al. 2017, A&A, 601, A16
Bautista, J. E., Busca, N. G., Guy, J., et al. 2017, ArXiv e-prints, arXiv:1702.00176
Becker, G. D., Bolton, J. S., & Lidz, A. 2015a, PASA, 32, e045
Becker, G. D., Bolton, J. S., Madau, P., et al. 2015b, MNRAS, 447, 3402
Becker, G. D., Rauch, M., & Sargent, W. L. W. 2007, ApJ, 662, 72
—. 2009, ApJ, 698, 1010
Becker, G. D., Sargent, W. L. W., Rauch, M., & Calverley, A. P. 2011, ApJ, 735, 93
Becker, G. D., Sargent, W. L. W., Rauch, M., & Carswell, R. F. 2012, ApJ, 744, 91
Becker, G. D., Sargent, W. L. W., Rauch, M., & Simcoe, R. A. 2006, ApJ, 640, 69
Becker, R. H., Fan, X., White, R. L., et al. 2001, AJ, 122, 2850
Bird, S., Rubin, K. H. R., Suresh, J., & Hernquist, L. 2016, MNRAS, 462, 307
Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Nature, 311, 517
Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93
Boera, E., Murphy, M. T., Becker, G. D., & Bolton, J. S. 2014, MNRAS, 441, 1916
Bohlin, R. C., Gordon, K. D., Rieke, G. H., et al. 2011, AJ, 141, 173
Boksenberg, A., & Sargent, W. L. W. 2015, ApJS, 218, 7
Bolton, J. S., Becker, G. D., Raskutti, S., et al. 2012, MNRAS, 419, 2880
Bolton, J. S., & Haehnelt, M. G. 2007a, MNRAS, 381, L35
—. 2007b, MNRAS, 382, 325
—. 2013, MNRAS, 429, 1695
Bolton, J. S., Haehnelt, M. G., Warren, S. J., et al. 2011, MNRAS, 416, L70
Bolton, J. S., Puchwein, E., Sijacki, D., et al. 2017, MNRAS, 464, 897
Bond, J. R., Kofman, L., & Pogosyan, D. 1996, Nature, 380, 603
Borthakur, S., Heckman, T. M., Leitherer, C., & Overzier, R. A. 2014, Science, 346, 216
Bosman, S. E. I., & Becker, G. D. 2015, MNRAS, 452, 1105
Bosman, S. E. I., Becker, G. D., Haehnelt, M. G., et al. 2017, MNRAS, 470, 1919
Bouché, N., Hohensee, W., Vargas, R., et al. 2012, MNRAS, 426, 801
Bouché, N., Murphy, M. T., Péroux, C., et al. 2007, ApJ, 669, L5

Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Bouwens, R. J., Oesch, P. A., Illingworth, G. D., Ellis, R. S., & Stefanon, M. 2016, ArXiv e-prints, arXiv:1610.00283
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011, ApJ, 737, 90
—. 2012, ApJ, 752, L5
Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645
Boylan-Kolchin, M., Bullock, J. S., & Kaplinghat, M. 2011, MNRAS, 415, L40
—. 2012, MNRAS, 422, 1203
Bradshaw, E. J., Almaini, O., Hartley, W. G., et al. 2013, MNRAS, 433, 194
Bromm, V., & Yoshida, N. 2011, ARA&A, 49, 373
Bullock, J. S., Kolatt, T. S., Sigad, Y., et al. 2001, MNRAS, 321, 559
Burbidge, E. M., Lynds, C. R., & Burbidge, G. R. 1966, ApJ, 144, 447
Burchett, J. N., Tripp, T. M., Bordoloi, R., et al. 2016, ApJ, 832, 124
Buzzoni, B., Delabre, B., Dekker, H., et al. 1984, The Messenger, 38, 9
Calura, F., Tescari, E., D'Odorico, V., et al. 2012, MNRAS, 422, 3019
Calverley, A. P., Becker, G. D., Haehnelt, M. G., & Bolton, J. S. 2011, MNRAS, 412, 2543
Caputi, K. I., Cirasuolo, M., Dunlop, J. S., et al. 2011, MNRAS, 413, 162
Carilli, C. L., Wang, R., Fan, X., et al. 2010, ApJ, 714, 834
Carnall, A. C., Shanks, T., Chehade, B., et al. 2015, MNRAS, 451, L16
Carswell, R. F., & Webb, J. K. 2014, VPFIT: Voigt profile fitting program, Astrophysics Source Code Library, ascl:1408.015
Caruana, J., Bunker, A. J., Wilkins, S. M., et al. 2012, MNRAS, 427, 3055
Cassata, P., Le Fèvre, O., Garilli, B., et al. 2011, A&A, 525, A143
Cassata, P., Tasca, L. A. M., Le Fèvre, O., et al. 2015, A&A, 573, A24
Castellano, M., Yue, B., Ferrara, A., et al. 2016, ApJ, 823, L40
Cen, R., & Kimm, T. 2015, ApJ, 801, L25
Chardin, J., Haehnelt, M. G., Aubert, D., & Puchwein, E. 2015, MNRAS, 453, 2943
Chardin, J., Haehnelt, M. G., Bosman, S. E. I., & Puchwein, E. 2017a, ArXiv e-prints, arXiv:1707.03841
Chardin, J., Puchwein, E., & Haehnelt, M. G. 2017b, MNRAS, 465, 3429
Chen, H.-W., Wild, V., Tinker, J. L., et al. 2010a, ApJ, 724, L176
Chen, S.-F. S., Simcoe, R. A., Torrey, P., et al. 2016, ArXiv e-prints, arXiv:1612.02829
Chen, Y.-M., Tremonti, C. A., Heckman, T. M., et al. 2010b, AJ, 140, 445
Choi, J.-H., & Nagamine, K. 2011, MNRAS, 410, 2579
Chornock, R., Berger, E., Fox, D. B., et al. 2013, ApJ, 774, 26
Choudhury, T. R., Haehnelt, M. G., & Regan, J. 2009, MNRAS, 394, 960
Choudhury, T. R., Puchwein, E., Haehnelt, M. G., & Bolton, J. S. 2015, MNRAS, 452, 261
Churchill, C. W., Nielsen, N. M., Kacprzak, G. G., & Trujillo-Gomez, S. 2013a, ApJ, 763, L42
Churchill, C. W., Rigby, J. R., Charlton, J. C., & Vogt, S. S. 1999, ApJS, 120, 51
Churchill, C. W., Trujillo-Gomez, S., Nielsen, N. M., & Kacprzak, G. G. 2013b, ApJ, 779, 87
Ciardi, B., Bolton, J. S., Maselli, A., & Graziani, L. 2012, MNRAS, 423, 558
Codoreanu, A., Ryan-Weber, E. V., Crighton, N. H. M., et al. 2017, MNRAS, 472, 1023
Coil, A. L., Weiner, B. J., Holz, D. E., et al. 2011, ApJ, 743, 46
Compostella, M., Cantalupo, S., & Porciani, C. 2013, MNRAS, 435, 3169
Cool, R. J., Kochanek, C. S., Eisenstein, D. J., et al. 2006, AJ, 132, 823
Cowie, L. L., Barger, A. J., & Hu, E. M. 2011, ApJ, 738, 136

Crighton, N. H. M., Murphy, M. T., Prochaska, J. X., et al. 2015, MNRAS, 452, 217
Croft, R. A. C. 1998, in Eighteenth Texas Symposium on Relativistic Astrophysics, ed. A. V. Olinto, J. A. Frieman, & D. N. Schramm, 664
Croft, R. A. C. 2004, ApJ, 610, 642
Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11
Curtis, M., & Sijacki, D. 2016, MNRAS, 463, 63
Curtis-Lake, E., McLure, R. J., Pearce, H. J., et al. 2012, MNRAS, 422, 1425
D'Aloisio, A., McQuinn, M., & Trac, H. 2015, ApJ, 813, L38
Davé, R., Katz, N., Oppenheimer, B. D., Kollmeier, J. A., & Weinberg, D. H. 2013, MNRAS, 434, 2645
Davé, R., Oppenheimer, B. D., & Finlator, K. 2011, MNRAS, 415, 11
Davies, F. B., & Furlanetto, S. R. 2016, MNRAS, 460, 1328
Davies, F. B., Hennawi, J. F., Eilers, A.-C., & Lukić, Z. 2017, ArXiv e-prints, arXiv:1703.10174
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
Dawson, S., Rhoads, J. E., Malhotra, S., et al. 2007, ApJ, 671, 1227
de Barros, S., Vanzella, E., Amorín, R., et al. 2016, A&A, 585, A51
De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2
De Rosa, G., Venemans, B. P., Decarli, R., et al. 2014, ApJ, 790, 145
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785
Dekel, A., & Silk, J. 1986, ApJ, 303, 39
Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, in Proc. SPIE, Vol. 4008, Optical and IR Telescope Instrumentation and Detectors, ed. M. Iye & A. F. Moorwood, 534–545
Delhaize, J., Meyer, M. J., Staveley-Smith, L., & Boyle, B. J. 2013, MNRAS, 433, 1398
Dijkstra, M., Mesinger, A., & Wyithe, J. S. B. 2011, MNRAS, 414, 2139
Dodelson, S. 2003, Modern cosmology
D'Odorico, V., Calura, F., Cristiani, S., & Viel, M. 2010, MNRAS, 401, 2715
D'Odorico, V., Cristiani, S., Romano, D., Granato, G. L., & Danese, L. 2004, MNRAS, 351, 976
D'Odorico, V., Cupani, G., Cristiani, S., et al. 2013, MNRAS, 435, 1198
Dubois, Y., Devriendt, J., Teyssier, R., & Slyz, A. 2011, MNRAS, 417, 1853
Dunlop, J. S. 2013, in Astrophysics and Space Science Library, Vol. 396, The First Galaxies, ed. T. Wiklind, B. Mobasher, & V. Bromm, 223
Eilers, A.-C., Davies, F. B., Hennawi, J. F., et al. 2017, ApJ, 840, 24
Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, MNRAS, 414, 888
Fan, X., Narayanan, V. K., Strauss, M. A., et al. 2002, AJ, 123, 1247
Fan, X., Strauss, M. A., Schneider, D. P., et al. 1999, AJ, 118, 1
Fan, X., White, R. L., Davis, M., et al. 2000, AJ, 120, 1167
Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001, AJ, 122, 2833
Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, AJ, 125, 1649
Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, AJ, 128, 515
Fan, X., Strauss, M. A., Becker, R. H., et al. 2006, AJ, 132, 117
Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., & Hernquist, L. 2009, ApJ, 703, 1416
Feng, L., Bi, H., Liu, J., & Fang, L.-Z. 2008, MNRAS, 383, 1459

Finkelstein, S. L., Papovich, C., Ryan, R. E., et al. 2012, ApJ, 758, 93
Finkelstein, S. L., Papovich, C., Dickinson, M., et al. 2013, Nature, 502, 524
Finlator, K., Muñoz, J. A., Oppenheimer, B. D., et al. 2013, MNRAS, 436, 1818
Finlator, K., Oppenheimer, B. D., Davé, R., et al. 2016, MNRAS, 459, 2299
Finlator, K., Thompson, R., Huang, S., et al. 2015, MNRAS, 447, 2526
Finlator, K., Prescott, M. K. M., Oppenheimer, B. D., et al. 2017, MNRAS, 464, 1633
Fiore, F., Puccetti, S., Grazian, A., et al. 2012, A&A, 537, A16
Ford, A. B., Davé, R., Oppenheimer, B. D., et al. 2014, MNRAS, 444, 1260
Forero-Romero, J. E., Yepes, G., Gottlöber, S., & Prada, F. 2012, MNRAS, 419, 952
Frieman, J. A., Turner, M. S., & Huterer, D. 2008, ARA&A, 46, 385
Fukugita, M., Ichikawa, T., Gunn, J. E., et al. 1996, AJ, 111, 1748
Furlanetto, S. R., & Oh, S. P. 2005, MNRAS, 363, 1031
—. 2008, ApJ, 681, 1
Furlanetto, S. R., Zaldarriaga, M., & Hernquist, L. 2006, MNRAS, 365, 1012
Gallerani, S., Choudhury, T. R., & Ferrara, A. 2006, MNRAS, 370, 1401
Gallerani, S., Ferrara, A., Fan, X., & Choudhury, T. R. 2008, MNRAS, 386, 359
García, L. A., Tescari, E., Ryan-Weber, E. V., & Wyithe, J. S. B. 2017, MNRAS, 469, L53
Garzilli, A., Bolton, J. S., Kim, T.-S., Leach, S., & Viel, M. 2012, MNRAS, 424, 1723
Gehrels, N. 1986, ApJ, 303, 336
Giallongo, E., Grazian, A., Fiore, F., et al. 2015, A&A, 578, A83
Glikman, E., Djorgovski, S. G., Stern, D., et al. 2011, ApJ, 728, L26
Gnedin, N. Y., Becker, G. D., & Fan, X. 2016, ArXiv e-prints, arXiv:1605.03183
Gnedin, N. Y., & Fan, X. 2006, ApJ, 648, 1
González, V., Labbé, I., Bouwens, R. J., et al. 2011, ApJ, 735, L34
Goto, T. 2006, MNRAS, 371, 769
Grazian, A., Giallongo, E., Gerbasi, R., et al. 2016, A&A, 585, A48
Greig, B., & Mesinger, A. 2017, MNRAS, 465, 4838
Greig, B., Mesinger, A., Haiman, Z., & Simcoe, R. A. 2016, MNRAS, arXiv:1606.00441
—. 2017, MNRAS, 466, 4239
Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, ApJ, 698, 198
Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633
Haardt, F., & Madau, P. 2012, ApJ, 746, 125
Heckman, T. M., Borthakur, S., Overzier, R., et al. 2011, ApJ, 730, 5
Henry, A., Scarlata, C., Martin, C. L., & Erb, D. 2015, ApJ, 809, 19
Hewett, P. C., & Wild, V. 2010, VizieR Online Data Catalog, 740
Heyl, J. S., Hernquist, L., & Spergel, D. N. 1994, ApJ, 427, 165
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hoag, A., Bradač, M., Trenti, M., et al. 2017, Nature Astronomy, 1, 0091
Hoffman, L., Cox, T. J., Dutta, S., & Hernquist, L. 2010, ApJ, 723, 818
Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129
Hook, I. M., Jørgensen, I., Allington-Smith, J. R., et al. 2004, PASP, 116, 425
Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581
Hopkins, P. F., Quataert, E., & Murray, N. 2012, MNRAS, 421, 3522
Hopkins, P. F., Bundy, K., Croton, D., et al. 2010, ApJ, 715, 202
Horne, K. 1986, PASP, 98, 609
Howarth, I. 2010, Astrophysical Processes: Nebulae to Stars; Part 4, Interstellar Absorption

Lines, online lecture notes
Hui, L., & Gnedin, N. Y. 1997, MNRAS, 292, 27
Hummels, C. B., Bryan, G. L., Smith, B. D., & Turk, M. J. 2013, MNRAS, 430, 1548
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Iliev, I. T., Mellema, G., Pen, U.-L., et al. 2006, MNRAS, 369, 1625
Inoue, A. K., Shimizu, I., Iwata, I., & Tanaka, M. 2014, MNRAS, 442, 1805
Iršič, V., Viel, M., Haehnelt, M. G., et al. 2017, ArXiv e-prints, arXiv:1702.01764
Izotov, Y. I., Orlitová, I., Schaerer, D., et al. 2016, Nature, 529, 178
Jaacks, J., Thompson, R., & Nagamine, K. 2013, ApJ, 766, 94
Jackson, J. C. 1972, MNRAS, 156, 1P
James, B. L., Auger, M., Aloisi, A., Calzetti, D., & Kewley, L. 2016, ApJ, 816, 40
Jaskot, A. E., & Oey, M. S. 2013, ApJ, 766, 91
Jeeson-Daniel, A., Ciardi, B., Maio, U., et al. 2012, MNRAS, 424, 2193
Jensen, H., Laursen, P., Mellema, G., et al. 2013, MNRAS, 428, 1366
Jiang, L., Fan, X., Vestergaard, M., et al. 2007, AJ, 134, 1150
Jiang, L., McGreer, I. D., Fan, X., et al. 2015, AJ, 149, 188
Jiang, L., Fan, X., Annis, J., et al. 2008, AJ, 135, 1057
Jiang, L., Fan, X., Bian, F., et al. 2009, AJ, 138, 305
Jiang, L., Finlator, K., Cohen, S. H., et al. 2016a, ApJ, 816, 16
Jiang, L., McGreer, I. D., Fan, X., et al. 2016b, ApJ, 833, 222
Jones, G. C., Willott, C. J., Carilli, C. L., et al. 2017, ArXiv e-prints, arXiv:1706.09968
Kacprzak, G. G., Churchill, C. W., Barton, E. J., & Cooke, J. 2011, ApJ, 733, 105
Kaiser, N. 1987, MNRAS, 227, 1
Kaiser, N., Burgett, W., Chambers, K., et al. 2010, in Proc. SPIE, Vol. 7733, Ground-based and Airborne Telescopes III, 77330E
Karman, W., Caputi, K. I., Trager, S. C., Almaini, O., & Cirasuolo, M. 2014, A&A, 565, A5
Kashikawa, N., Aoki, K., Asai, R., et al. 2002, PASJ, 54, 819
Kashikawa, N., Shimasaku, K., Matsuda, Y., et al. 2011, ApJ, 734, 119
Kauffmann, G. 1996, MNRAS, 281, 487
Kauffmann, G., & White, S. D. M. 1993, MNRAS, 261, doi:10.1093/mnras/261.4.921
Kausch, W., Noll, S., Smette, A., et al. 2015, A&A, 576, A78
Keating, L. C., Haehnelt, M. G., Becker, G. D., & Bolton, J. S. 2014, MNRAS, 438, 1820
Keating, L. C., Haehnelt, M. G., Cantalupo, S., & Puchwein, E. 2015, MNRAS, 454, 681
Keating, L. C., Puchwein, E., & Haehnelt, M. G. 2017, ArXiv e-prints, arXiv:1709.05351
Keating, L. C., Puchwein, E., Haehnelt, M. G., Bird, S., & Bolton, J. S. 2016, MNRAS, 461, 606
Kelson, D. D. 2003, PASP, 115, 688
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Khaire, V., & Srianand, R. 2015, MNRAS, 451, L30
Khochfar, S., & Silk, J. 2009, ApJ, 700, L21
Kim, Y., Im, M., Jeon, Y., et al. 2015, ApJ, 813, L35
Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, ApJ, 522, 82
KOA. 2017, W. M. Keck Observatory Archive
Kolmogorov, A. 1933, Inst. Ital. Attuari, Giorn., 4, 83
Kornei, K. A., Shapley, A. E., Erb, D. K., et al. 2010, ApJ, 711, 693
Krajnović, D., Emsellem, E., Cappellari, M., et al. 2011, MNRAS, 414, 2923

Lacey, C., & Cole, S. 1993, MNRAS, 262, 627
Laursen, P., Razoumov, A. O., & Sommer-Larsen, J. 2009, ApJ, 696, 853
Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
Lidz, A., McQuinn, M., Zaldarriaga, M., Hernquist, L., & Dutta, S. 2007, ApJ, 670, 39
Lidz, A., Oh, S. P., & Furlanetto, S. R. 2006, ApJ, 639, L47
Liu, C., Mutch, S. J., Angel, P. W., et al. 2016, MNRAS, 462, 235
Livermore, R. C., Finkelstein, S. L., & Lotz, J. M. 2017, ApJ, 835, 113
Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, ApJ, 837, 97
Lovell, M. R., Eke, V. R., Frenk, C. S., & Jenkins, A. 2011, MNRAS, 413, 3013
Lundgren, B. F., Brunner, R. J., York, D. G., et al. 2009, ApJ, 698, 819
Lyman, T. 1906, ApJ, 23, 181
Lynds, C. R., & Stockton, A. N. 1966, ApJ, 144, 446
Macciò, A. V., Kang, X., & Moore, B. 2009, ApJ, 692, L109
MacLeod, C. L., Ross, N. P., Lawrence, A., et al. 2016, MNRAS, 457, 389
Madau, P., & Haardt, F. 2015, ApJ, 813, L8
Madau, P., Haardt, F., & Rees, M. J. 1999, ApJ, 514, 648
Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, AJ, 115, 2285
Marshall, J. L., Burles, S., Thompson, I. B., et al. 2008, in Proc. SPIE, Vol. 7014, Ground-based and Airborne Instrumentation for Astronomy II, 701454
Martin, A. M., Papastergis, E., Giovanelli, R., et al. 2010, ApJ, 723, 1359
Martin, C. L. 2005, ApJ, 621, 227
—. 2006, ApJ, 647, 222
Mas-Ribas, L., Hennawi, J. F., Dijkstra, M., et al. 2017, ArXiv e-prints, arXiv:1707.07713
Maselli, A., Ferrara, A., & Gallerani, S. 2009, MNRAS, 395, 1925
Maselli, A., Gallerani, S., Ferrara, A., & Choudhury, T. R. 2007, MNRAS, 376, L34
Massey, P., & Hanson, M. M. 2013, Astronomical Spectroscopy, ed. T. D. Oswalt & H. E. Bond, 35
Matejek, M. S., & Simcoe, R. A. 2012, ApJ, 761, 112
Matejek, M. S., Simcoe, R. A., Cooksey, K. L., & Seyffert, E. N. 2013, ApJ, 764, 9
Mathes, N. L., Churchill, C. W., & Murphy, M. T. 2017, ArXiv e-prints, arXiv:1701.05624
Matsuoka, Y., Onoue, M., Kashikawa, N., et al. 2016, ApJ, 828, 26
—. 2017, ArXiv e-prints, arXiv:1704.05854
Matthee, J., Sobral, D., Oteo, I., et al. 2016, MNRAS, 458, 449
Matzeu, G. A., Reeves, J. N., Braito, V., et al. 2017, ArXiv e-prints, arXiv:1708.03546
Mazzucchelli, C., Bañados, E., Venemans, B. P., et al. 2017, ArXiv e-prints, arXiv:1710.01251
McDonald, P., Seljak, U., Cen, R., Bode, P., & Ostriker, J. P. 2005, MNRAS, 360, 1471
McGreer, I. D., Mesinger, A., & D'Odorico, V. 2015, MNRAS, 447, 499
McGreer, I. D., Mesinger, A., & Fan, X. 2011, MNRAS, 415, 3237
McLean, I. S., Steidel, C. C., Epps, H., et al. 2010, in Proc. SPIE, Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III, 77351E–77351E–12
McMahon, R. G., Banerji, M., Gonzalez, E., et al. 2013, The Messenger, 154, 35
McQuinn, M., Hernquist, L., Lidz, A., & Zaldarriaga, M. 2011, MNRAS, 415, 977
McQuinn, M., Lidz, A., Zahn, O., et al. 2007, MNRAS, 377, 1043
McQuinn, M., Lidz, A., Zaldarriaga, M., et al. 2009, ApJ, 694, 842
Meiksin, A. 2005, MNRAS, 356, 596
Ménard, B., Wild, V., Nestor, D., et al. 2011, MNRAS, 417, 801

Mesinger, A. 2010, MNRAS, 407, 1328
Mesinger, A., Aykutalp, A., Vanzella, E., et al. 2015, MNRAS, 446, 566
Mesinger, A., & Haiman, Z. 2004, ApJ, 611, L69
—. 2007, ApJ, 660, 923
Miralda-Escudé, J. 1998, ApJ, 501, 15
—. 2000, ApJ, 528, L1
Misawa, T., Charlton, J. C., Eracleous, M., et al. 2007, ApJS, 171, 1
Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84460Z
Moore, B., Ghigna, S., Governato, F., et al. 1999, ApJ, 524, L19
Morganson, E., De Rosa, G., Decarli, R., et al. 2012, AJ, 143, 142
Mortlock, D. J., Patel, M., Warren, S. J., et al. 2009, A&A, 505, 97
Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616
Mostardi, R. E., Shapley, A. E., Steidel, C. C., et al. 2015, ApJ, 810, 107
Moster, B. P., Macciò, A. V., Somerville, R. S., Naab, T., & Cox, T. J. 2011, MNRAS, 415, 3750
Mutch, S. J., Croton, D. J., & Poole, G. B. 2013, MNRAS, 435, 2445
Naab, T., Oser, L., Emsellem, E., et al. 2014, MNRAS, 444, 3357
Nakajima, K., & Ouchi, M. 2014, MNRAS, 442, 900
Nestor, D. B., Turnshek, D. A., & Rao, S. M. 2005, ApJ, 628, 637
Nilsson, K. K., Tapken, C., Møller, P., et al. 2009, A&A, 498, 13
Noll, S., Kausch, W., Kimeswenger, S., et al. 2014, A&A, 567, A25
Noterdaeme, P., Petitjean, P., Carithers, W. C., et al. 2012, A&A, 547, L1
Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, MNRAS, 463, 1462
Oh, S. P., & Furlanetto, S. R. 2005, ApJ, 620, L9
Okamoto, T., Frenk, C. S., Jenkins, A., & Theuns, T. 2010, MNRAS, 406, 208
Oke, J. B., & Gunn, J. E. 1982, PASP, 94, 586
Ono, Y., Ouchi, M., Mobasher, B., et al. 2012, ApJ, 744, 83
Oppenheimer, B. D., & Davé, R. 2006, MNRAS, 373, 1265
Oppenheimer, B. D., Davé, R., & Finlator, K. 2009, MNRAS, 396, 729
Oppenheimer, B. D., Davé, R., Katz, N., Kollmeier, J. A., & Weinberg, D. H. 2012, MNRAS, 420, 829
Oppenheimer, B. D., Davé, R., Kereš, D., et al. 2010, MNRAS, 406, 2325
Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, ApJ, 725, 2312
Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2010, ApJ, 723, 869
Oyarzún, G. A., Blanc, G. A., González, V., et al. 2016, ApJ, 821, L14
Pallottini, A., Ferrara, A., Gallerani, S., Salvadori, S., & D'Odorico, V. 2014, MNRAS, 440, 2498
Pâris, I., Petitjean, P., Rollinde, E., et al. 2011, A&A, 530, A50
Pâris, I., Petitjean, P., Ross, N. P., et al. 2017, A&A, 597, A79
Parsa, S., Dunlop, J. S., & McLure, R. J. 2017, ArXiv e-prints, arXiv:1704.07750
Pentericci, L., Fontana, A., Vanzella, E., et al. 2011, ApJ, 743, 132
Pentericci, L., Vanzella, E., Fontana, A., et al. 2014, ApJ, 793, 113
Péroux, C., Dessauges-Zavadsky, M., D'Odorico, S., Sun Kim, T., & McMahon, R. G. 2005, MNRAS, 363, 479

Pettini, M. 2010, Stellar Atmospheres II: The Analysis of Absorption Lines, online lecture notes
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A13
Planck Collaboration, Adam, R., Aghanim, N., et al. 2016b, A&A, 596, A108
Pogge, R. W., Atwood, B., O'Brien, T. P., et al. 2012, in Proc. SPIE, Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV, 84460G
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Proga, D., Stone, J. M., & Kallman, T. R. 2000, ApJ, 543, 686
Puchwein, E., Pfrommer, C., Springel, V., Broderick, A. E., & Chang, P. 2012, MNRAS, 423, 149
Puchwein, E., & Springel, V. 2013, MNRAS, 428, 2966
Qin, Y., Mutch, S. J., Poole, G. B., et al. 2017, ArXiv e-prints, arXiv:1703.04895
Rafelski, M., Neeleman, M., Fumagalli, M., Wolfe, A. M., & Prochaska, J. X. 2014, ApJ, 782, L29
Rahmati, A., Schaye, J., Crain, R. A., et al. 2016, MNRAS, 459, 310
Rauch, M. 1998, ARA&A, 36, 267
Reed, S. L., McMahon, R. G., Banerji, M., et al. 2015, MNRAS, 454, 3952
Reed, S. L., McMahon, R. G., Martini, P., et al. 2017, MNRAS, 468, 4702
Richards, G. T., Vanden Berk, D. E., Reichard, T. A., et al. 2002, AJ, 124, 1
Richards, G. T., Kruczek, N. E., Gallagher, S. C., et al. 2011, AJ, 141, 167
Roberts-Borsani, G. W., Bouwens, R. J., Oesch, P. A., et al. 2016, ApJ, 823, 143
Robertson, B. E., Ellis, R. S., Furlanetto, S. R., & Dunlop, J. S. 2015, ApJ, 802, L19
Robertson, B. E., Furlanetto, S. R., Schneider, E., et al. 2013, ApJ, 768, 71
Rorai, A., Hennawi, J. F., & White, M. 2013, ApJ, 775, 81
Rorai, A., Becker, G. D., Haehnelt, M. G., et al. 2017, MNRAS, 466, 2690
Rupke, D. S., Veilleux, S., & Sanders, D. B. 2005, ApJS, 160, 115
Ryan-Weber, E. V., Pettini, M., & Madau, P. 2006, MNRAS, 371, L78
Ryan-Weber, E. V., Pettini, M., Madau, P., & Zych, B. J. 2009, MNRAS, 395, 1476
Savage, B. D., & Sembach, K. R. 1991, ApJ, 379, 245
Schaye, J., Aguirre, A., Kim, T.-S., et al. 2003, ApJ, 596, 768
Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536
Schenker, M. A., Stark, D. P., Ellis, R. S., et al. 2012, ApJ, 744, 179
Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, ApJ, 768, 196
Schmidt, G. D., Weymann, R. J., & Foltz, C. B. 1989, PASP, 101, 713
Schroeder, J., Mesinger, A., & Haiman, Z. 2013, MNRAS, 428, 3058
Shang, Z., Wills, B. J., Wills, D., & Brotherton, M. S. 2007, AJ, 134, 294
Sheinis, A. I., Bolte, M., Epps, H. W., et al. 2002, PASP, 114, 851
Shen, S., Madau, P., Aguirre, A., et al. 2012, ApJ, 760, 50
Shen, S., Madau, P., Conroy, C., Governato, F., & Mayer, L. 2014, ApJ, 792, 99
Shull, J. M., France, K., Danforth, C. W., Smith, B., & Tumlinson, J. 2010, ApJ, 722, 1312
Siana, B., Shapley, A. E., Kulas, K. R., et al. 2015, ApJ, 804, 17
Simcoe, R. A. 2006, ApJ, 653, 977
Simcoe, R. A., Sargent, W. L. W., Rauch, M., & Becker, G. 2006, ApJ, 637, 648
Simcoe, R. A., Sullivan, P. W., Cooksey, K. L., et al. 2012, Nature, 492, 79
Simcoe, R. A., Cooksey, K. L., Matejek, M., et al. 2011, ApJ, 743, 21
Simon, L. E., Hamann, F., & Pettini, M. 2012, in Astronomical Society of the Pacific Con-

ference Series, Vol. 460, AGN Winds in Charleston, ed. G. Chartas, F. Hamann, & K. M. Leighly, 52
Smette, A., Sana, H., Noll, S., et al. 2015, A&A, 576, A77
Smith, M. C., Sijacki, D., & Shen, S. 2017, ArXiv e-prints, arXiv:1709.03515
Songaila, A. 2004, AJ, 127, 2598
Songaila, A., & Cowie, L. L. 2010, ApJ, 721, 1448
Springel, V. 2005, MNRAS, 364, 1105
Stark, D. P., Ellis, R. S., Bunker, A., et al. 2009, ApJ, 697, 1493
Stark, D. P., Ellis, R. S., Chiu, K., Ouchi, M., & Bunker, A. 2010, MNRAS, 408, 1628
Stark, D. P., Ellis, R. S., Charlot, S., et al. 2017, MNRAS, 464, 469
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Steidel, C. C., Giavalisco, M., Dickinson, M., & Adelberger, K. L. 1996, AJ, 112, 352
Suresh, J., Bird, S., Vogelsberger, M., et al. 2015, MNRAS, 448, 895
Suresh, J., Rubin, K. H. R., Kannan, R., et al. 2017, MNRAS, 465, 2966
Suzuki, N., Tytler, D., Kirkman, D., O'Meara, J. M., & Lubin, D. 2005, ApJ, 618, 592
Syphers, D., Anderson, S. F., Zheng, W., et al. 2009, ApJS, 185, 20
Tanvir, N. R., Fox, D. B., Levan, A. J., et al. 2009, Nature, 461, 1254
Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. 2006, Phys. Rev. D, 74, 123507
Tescari, E., Viel, M., D'Odorico, V., et al. 2011, MNRAS, 411, 826
The Dark Energy Survey Collaboration. 2005, ArXiv Astrophysics e-prints, astro-ph/0510346
Theuns, T., Leonard, A., Efstathiou, G., Pearce, F. R., & Thomas, P. A. 1998, MNRAS, 301, 478
Theuns, T., Viel, M., Kay, S., et al. 2002, ApJ, 578, L5
Tilvi, V., Papovich, C., Finkelstein, S. L., et al. 2014, ApJ, 794, 5
Totani, T. 2013, PASJ, 65, L12
Trac, H., & Cen, R. 2007, ApJ, 671, 1
Trainor, R. F., Steidel, C. C., Strom, A. L., & Rudie, G. C. 2015, ApJ, 809, 89
Treu, T., Schmidt, K. B., Trenti, M., Bradley, L. D., & Stiavelli, M. 2013, ApJ, 775, L29
Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, Science, 334, 948
Turner, M. L., Schaye, J., Steidel, C. C., Rudie, G. C., & Strom, A. L. 2014, MNRAS, 445, 794
Ueda, Y., Akiyama, M., Ohta, K., & Miyaji, T. 2003, ApJ, 598, 886
Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549
Vanzella, E., Siana, B., Cristiani, S., & Nonino, M. 2010, MNRAS, 404, 1672
Vanzella, E., Guo, Y., Giavalisco, M., et al. 2012, ApJ, 751, 70
Vanzella, E., De Barros, S., Cupani, G., et al. 2016, ApJ, 821, L27
Venemans, B. P., McMahon, R. G., Warren, S. J., et al. 2007, MNRAS, 376, L76
Venemans, B. P., McMahon, R. G., Walter, F., et al. 2012, ApJ, 751, L25
Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, ApJ, 779, 24
Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, MNRAS, 453, 2259
Verner, D. A., & Ferland, G. J. 1996, ApJS, 103, 467
Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, A&A, 536, A105
Viel, M., Becker, G. D., Bolton, J. S., & Haehnelt, M. G. 2013, Phys. Rev. D, 88, 043502
Viel, M., Bolton, J. S., & Haehnelt, M. G. 2009, MNRAS, 399, L39
Viel, M., Lesgourgues, J., Haehnelt, M. G., Matarrese, S., & Riotto, A. 2005, Phys. Rev. D, 71, 063534

Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, in Proc. SPIE, Vol. 2198, Instrumentation in Astronomy VIII, ed. D. L. Crawford & E. R. Craine, 362
Wang, F., Wu, X.-B., Fan, X., et al. 2016, ApJ, 819, 24
Weinberg, S. 2008, Cosmology (Oxford University Press)
Weiner, B. J., Coil, A. L., Prochaska, J. X., et al. 2009, ApJ, 692, 187
Weymann, R. J., Carswell, R. F., & Smith, M. G. 1981, ARA&A, 19, 41
Weymann, R. J., Williams, R. E., Peterson, B. M., & Turnshek, D. A. 1979, ApJ, 234, 33
White, R. L., Becker, R. H., Fan, X., & Strauss, M. A. 2003, AJ, 126, 1
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
Wiersma, R. P. C., Schaye, J., Dalla Vecchia, C., et al. 2010, MNRAS, 409, 132
Wilczynska, M. R., Webb, J. K., King, J. A., et al. 2015, MNRAS, 454, 3082
Willott, C. J., Delorme, P., Omont, A., et al. 2007, AJ, 134, 2435
Willott, C. J., Delorme, P., Reylé, C., et al. 2009, AJ, 137, 3541
—. 2010, AJ, 139, 906
Wise, J. H., Demchenko, V. G., Halicek, M. T., et al. 2014, MNRAS, 442, 2560
Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2005, ARA&A, 43, 861
Wolfe, A. M., Turnshek, D. A., Smith, H. E., & Cohen, R. D. 1986, ApJS, 61, 249
Wootten, A., & Thompson, A. R. 2009, IEEE Proceedings, 97, 1463
Worseck, G., Prochaska, J. X., Hennawi, J. F., & McQuinn, M. 2016, ApJ, 825, 144
Wu, J., Charlton, J. C., Misawa, T., Eracleous, M., & Ganguly, R. 2010, ApJ, 722, 997
Wu, X.-B., Wang, F., Fan, X., et al. 2015, Nature, 518, 512
Wyithe, J. S. B., & Bolton, J. S. 2011, MNRAS, 412, 1926
Wyithe, J. S. B., Bolton, J. S., & Haehnelt, M. G. 2008, MNRAS, 383, 691
Wyithe, J. S. B., Loeb, A., & Oesch, P. A. 2014, MNRAS, 439, 1326
Xu, Y., Yue, B., Su, M., Fan, Z., & Chen, X. 2014, ApJ, 781, 97
Yong, S. Y., Webster, R. L., King, A. L., et al. 2017, ArXiv e-prints, arXiv:1708.07237
York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579
Yue, B., Ferrara, A., & Xu, Y. 2016, MNRAS, 463, 1968
Zaldarriaga, M. 1997, Phys. Rev. D, 55, 1822
Zitrin, A., Ellis, R. S., Belli, S., & Stark, D. P. 2015, ApJ, 805, L7
Zwaan, M. A., Meyer, M. J., Staveley-Smith, L., & Webster, R. L. 2005, MNRAS, 359, L30